

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:14:19 ; Search time 24.3966 Seconds
(without alignments)
755.549 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLRYGFKC 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	607.5	31.5	353	1	P56542 mus musculus
2	604.5	31.4	364	1	O62855 sus scrofa
3	596.5	31.0	360	1	O00115 homo sapien
4	535	27.8	375	1	O17778 caenorhabdi
5	470.5	24.4	367	1	P34387 caenorhabdi
6	462.5	24.0	378	1	P34508 caenorhabdi
7	443	23.0	276	1	P56541 bos taurus
8	101.5	5.3	475	1	O9ubm7 homo sapien
9	94	4.9	1385	1	O45760 bacillus th
10	94	4.9	2165	1	P28887 human respi
11	93.5	4.9	582	1	P19252 pisum sativ
12	90	4.7	1276	1	P97250 cricetus
13	90	4.7	1445	1	P23470 homo sapien
14	88.5	4.6	1435	1	P47729 mycoplasma
15	88	4.6	3388	1	P12823 d genome po
16	87.5	4.5	175	1	P75143 mycoplasma
17	87.5	4.5	470	1	P33035 fremyella d
18	87.5	4.5	957	1	O8r553 rattus norv
19	87.5	4.5	1278	1	Q12770 homo sapien
20	87	4.5	900	1	P47534 mycoplasma
21	87	4.5	2109	1	P16379 vesicular s
22	86.5	4.5	360	1	P50773 human papil
23	86.5	4.5	1386	1	P75613 mycoplasma
24	85	4.4	867	1	P04025 simian retr
25	85	4.4	1442	1	Q05909 mus musculu
26	84.5	4.4	541	1	Q25861 plasmodium
27	84	4.4	675	1	P51813 homo sapien
28	84	4.4	956	1	O99j87 mus musculu
29	84	4.4	1025	1	P53708 homo sapien
30	84	4.4	5376	1	O88799 mus musculu
31	83.5	4.3	394	1	P49150 desmodus ro
32	83.5	4.3	524	1	P47635 mycoplasma
33	83.5	4.3	737	1	Q91225 mus musculu

34	83.5	4.3	1409	1	COPI DROME
35	83.5	4.3	1436	1	WRN_XENLA
36	83	4.3	2109	1	RRPL_VSVJH
37	82.5	4.3	980	1	POLG LIV
38	82	4.3	426	1	Y958 METJA
39	82	4.3	1473	1	OVOS CHICK
40	82	4.3	2351	1	FAB_HUMAN
41	81.5	4.2	496	1	POLG LIVK
42	81.5	4.2	496	1	POLG LIVNO
43	81.5	4.2	704	1	NEUL_PIG
44	81	4.2	428	1	SYH_BUCAP
45	81	4.2	559	1	TPA_MOUSE

ALIGNMENTS

RESULT 1					
DRN2_MOUSE					
ID	DRN2_MOUSE	STANDARD;	PRT;	353 AA.	
AC	P56542; O55053;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)				
DE	(Lysosomal DNase II)				
GN	DNASE2 OR DNASE2A OR DNL2.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98382521; PubMed=9714827;				
RA	Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;				
RT	"Molecular cloning and characterization of human and murine DNase II."				
RL	Gene 215:281-289 (1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20363736; PubMed=10903447;				
RA	Krieser R.J., Eastman A.;				
RT	"Deoxyribonuclease II: structure and chromosomal localization of the murine gene, and comparison with the genomic structure of the human and three C. elegans homologs."				
RL	Gene 252:153-162 (2000).				
CC	-1- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference for double-stranded DNA. Has a possible role in apoptosis.				
CC	-1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-phosphates and 3'-phosphooligonucleotide end-products.				
CC	-1- SUBCELLULAR LOCATION: Lysosomal.				
CC	-1- SIMILARITY: Belongs to the DNase II family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR	EMBL; AF045741; AAC35750.1; -				
DR	EMBL; AF190459; AAF20386.1; -				
DR	MGI; MGI:1329019; Dnase2a.				
DR	InterPro; IPR004947; DNase_II.				
DR	Pfam; PF03265; DNase II; 1.				
KW	Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Signal;				
KW	Apoptosis.				
FT	SIGNAL. 1 19				POTENTIAL.
FT	PROPEP 20 109				BY SIMILARITY.
FT	CHAIN 110 353				DEOXYRIBONUCLEASE II.
FT	CARBOHYD 71 71				N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 88 88				N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 353 AA; 38810 MW; F53CDD81CF6FCF4 CRC64;

Query Match 31.5%; Score 607.5; DB 1; Length 353;
Best Local Similarity 37.4%; Pred. No. 5.5e-45;
Matches 130; Conservative 59; Mismatches 146; Indels 13; Gaps 6;

QY 6 LRTVLISLFFALSGVLGTPEISCRNEYGEAVDWFIFYLKPKRTSKAS-EBAGLQYLYDST 65
Db 1 MATLSLLALLMW-PAEALSCGDSGQPDVFFVYKLPAPHSRDTPKGLTYKYMDO 59
66 RQTNWNSLYLINSTRSALGRTLOHLYDTHNSNDTAYLIYNDGVP--GSVNVSRQYGHAK 123
60 SDGWDQGVYINSEGA-GRSLQPLY--RKNSQLAFLLYNDQPKSSARDSTGCHGTHK 117
124 GLLVNRTQGGFWLHSHVPEKPPP---VHGVEYPTSGRRYGTGICITFGYSQFEEDFOLL 180
118 GVLLMDQGGFWLHSHVPRFPFPASSGAYTWPPNAQTGQTLCLVSLPFTQFARIGKQLT 177
181 VLQPNVYSCFIPSTFPHWKLIYMPRCANSSSIKIPVRYLAELHSAQGLNFVHFAKSSFT 240
178 YTYPLVYDHKLEGGFAQKLPDLETIKVQHLHBPWNSVILTSQAGATFQSFAPKFGK 237
241 DDIFGTMLAQKLTLLHQAQTQWKKQELPNSCSLPYHYVNIKSIGV--TSKSYFSSRDH 298
238 DDLVSGWLAELGNTLVQVFNQSGILPNSCSGAYQVLDVTOTGFPGRSLTFSATEDH 297
299 SKWCVSIKGSANRWTCIGDLNRSLSHQAALRGGFICTKNHYIYQAFHKL 346
298 SKWCVAPOQ---WACVGDMMNRKAEHTRGGGTCTQCTPFWKAFQSL 342

RESULT 2
DRN2_PIG
ID DRN2_HUMAN STANDARD; PRT; 364 AA.
AC O62855;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)
GN DNASE2 OR DNL2.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Wang C.C., Lu S.C., Chen H.L., Liao T.H.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=98321218; PubMed=9647784;
RA Shikawa D., Tanuma S.-I.;
RL "Cloning of cDNAs encoding porcine and human DNase II.";
RL Biochem. Biophys. Res. Commun. 247:864-869(1998).
CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference
CC for double-stranded DNA. Has a possible role in apoptosis.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
CC phosphates and 3'-phosphooligonucleotide end-products.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: Belongs to the DNase II family.
CC -!- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/HDAC/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001387; CAA04717.1; -.
DR EMBL; AF060221; AAC39263.1; -.
DR PIR; J0205; JE0205.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase_II; 1.
KW Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Apoptosis;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 108
FT CHAIN 109 364
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 40183 MW; 2B690BBFE33734D CRC64;

Query Match 31.4%; Score 604.5; DB 1; Length 364;
Best Local Similarity 37.9%; Pred. No. 1e-44;
Matches 136; Conservative 60; Mismatches 144; Indels 19; Gaps 9;

QY 6 LRTVLISLFFALSGV-LGTPEISCRNEYGEAVDWFIFYLKPKRTSKAS-EBAGLQYLYD 63
Db 1 MATLSPLLLAALLMWVPGT--LTCVSGQGPVDFVYKLPAPHSRPGDVAGSLYKYL 58
64 STROTWNKSLYLINSTRSALGRTLOHLYDTHNSNDTAYLIYNDGVP---GSVNVSRQY 120
59 EESGWRDAGSINSSTGALGRSLPLY--RNTSQLAFLLYNDQPKYRGS-QHSSNRG 114
121 HAKGLLVNRTQGGFWLHSHVPEKPPP---GYEYPTSGRRYGTGICITFGYSQFEEDIF 177
115 HTKGVLLDDQGGFWLHSHVFNFPFPPSSAAYSMPSPARTYGTQLICVSPFLTQPLNISR 174
178 QLLVLPNIYSCFIPSTFPHWKLIYMPRCANSSSIKIPVRYLAELHSAQGLNFVHFAKSS 237
175 QLTYPNVYDYKLEGGFAKFPFVLEEVVGHVLEQFPWNSVILTSKAGASFQSAKCG 234
238 FYTDIDFTGTAQKLTLLHQAQTQWKKQELPNSCSLPYHYVNIKSIGVTSKS--YFSGR 295
235 NFGDDLYSGWLAELGNTLVQVFNQSGILPNSCSGAYQVLDVTQIAFPGPAGNFNAT 294
296 QDHSKWCVSIKGSANRWTCIGDLNRSLSHQAALRGGFICTKNHYIYQAFHKLVLRYGFC 354
295 EDHSKWCVA---PERPWTCTVGDMMNRKFEHRRGGTLCALPLAKFPLVKAWPEPCB 350

RESULT 3
DRN2_HUMAN
ID DRN2_HUMAN STANDARD; PRT; 360 AA.
AC O00115; O43910;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)
GN DNASE2 OR DNL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112802; PubMed=9446563;
RA Yasuda T., Takeshita H., Iida R., Nakajima T., Hosomi O.,
RA Nakashima Y., Kishi K.;
RL "Molecular cloning of the cDNA encoding human deoxyribonuclease II.";
RL J. Biol. Chem. 273:2610-2616(1998).
RN [2]
RP SEQUENCE FROM N.A.
```


RX MEDLINE=20363736; PubMed=10903447;
 RA Krieser R.J., Eastman A.;
 RT "Deoxyribonuclease II: structure and chromosomal localization of the
 RT murine gene, and comparison with the genomic structure of the human
 RT and three C. elegans homologs.";
 RL Gene 252:155-162(2000).
 [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98382521; PubMed=9714827;
 RA Baker K.P., Barton W.F., Henzel W.J., Spencer S.A.;
 RT "Molecular cloning and characterization of human and murine DNase
 RT II.";
 RL Gene 215:281-289(1998).
 [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98321218; PubMed=9647784;
 RA Shiokawa D., Tanuma S.-I.;
 RT Cloning of cDNAs encoding porcine and human DNase II.";
 RL Biochem. Biophys. Res. Commun. 247:864-869(1998).
 [5]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99123683; PubMed=9924608;
 RA Yasuda T., Takeshita H., Iida R., Tsutsumi S., Nakajima T., Hosomi O.,
 RA Nakashima Y., Mori S., Kishi K.;
 RT "Structure and organization of the human deoxyribonuclease II (DNase
 RT II) gene.";
 RL Ann. Hum. Genet. 62:299-305(1998).
 [6]
 RN SEQUENCE FROM N.A.
 RP Lamerdin J.E., McCreedy P.M., Stilwagen S., Ramirez M., Carrano A.;
 RT "Characterization by genomic sequence analysis of a gene-rich 111 kb
 RT region of 19p13.2 containing the human DNA repair gene, RAD23A.";
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 [7]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [8]
 RN CARBOHYDRATE-LINKAGE SITE ASN-212.
 RP MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference
 CC for double-stranded DNA. Has a possible role in apoptosis.
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
 CC phosphates and 3'-phosphonucleotide end-products.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Glycosylated.
 CC -!- SIMILARITY: Belongs to the DNase II family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; AB004574; BAA28623.1; -
 DR ENBL; AF047016; AAC77366.1; -
 DR ENBL; AF045937; AAC35751.1; -
 DR ENBL; AF060222; AAC39852.1; -
 DR ENBL; AB008564; BAB55598.1; -
 DR ENBL; AB008559; BAB55598.1; JOINED.
 DR ENBL; AB008560; BAB55598.1; JOINED.
 DR ENBL; AB008561; BAB55598.1; JOINED.
 DR ENBL; AB008562; BAB55598.1; JOINED.
 DR ENBL; AB008563; BAB55598.1; JOINED.
 DR ENBL; AD000092; AAB51172.1; ALT_SEQ.
 DR ENBL; BC010419; AAH10419.1; -
 DR PIR; JE0206; JE0206.
 DR Genew; HGNC:2960; DNASE2.
 DR MIM; 126350; -
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0004531; F:deoxyribonuclease II activity; TAS.
 DR GO; GO:0003677; F:DNA binding; TAS.
 DR GO; GO:0006259; P:DNA metabolism; TAS.
 DR InterPro; IPR004947; DNase-II.
 DR Pfam; PF03285; DNase-II; 1.
 DR KW Hydroxylase; Endonuclease; Nuclease; Lysosome; Zymogen; Signal;
 KW Apoptosis; Glycoprotein; Polymorphism.
 FT SIGNAL 1 18
 FT PROPEP 19 106 BY SIMILARITY
 FT CHAIN 107 360 DEOXYRIBONUCLEASE II.
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 314 314 R -> L (in dbSNP:1061192).
 FT /FTID=VAR_012044.
 SQ SEQUENCE 360 AA; 39581 MW; DF1BFBFA8A9676EA CRC64;
 Query Match 31.0%; Score 596.5; DB 1; Length 360;
 Best Local Similarity 38.0%; Pred. No. 5e-44;
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;
 QY 17 LSGVLGTP--EISCNEYGEADVDFYKLPK-RTSKASEEAGLOLYLDSTQTNKSL 73
 Db 6 LAALLCPAGALTCYDGGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65
 QY 74 YLINSRSGALGTHLYDTHSTNDTAVLYND--GVFGSVNYSQYGHAKGLLVNRT 131
 Db 66 ALINSPGAVGKSLQPLY--RNTSQALFLYNDQPPQSKAQDSSNRGHTKGLLDHD 123
 QY 132 QGFWLIHVPKPPP---VHGYYFTSGRYYGTGICITFGYSQFEIDFOLLVQNIYS 188
 Db 124 GGFVLVHVPNPPPPASSAAAYSWPHSACTYGTLLCVSPFAQFSKMGKQLTYTPWYN 193
 QY 189 CIPSTFTHWKLIMPRMCANSSSLKIPVLYLAELHSAQGLNFVHFAKSSFYTDITGWI 248
 Db 184 YQLEGIFAQEFPPDLENVVGKHSVQEPWNSITLTQAGAVFQSFQKFKGDDLYSGWL 243
 QY 249 AQKLKTHLAQWQKKQELPNCISLPVYVNIKSGVTSKS--YESSQDHSKWCYSIK 306
 Db 244 AALGNLQVQVHHKTVTGILPNSCDIWOVLNVNIAFPFGPAGPFSNSTDHSKWCVSPK 303
 QY 307 GSANRWTCIGDNLRSIHLQALRGSGGFTCTKNHYTYQAFHKLKLYRGFC 353
 Db 304 GP---WTCVGDVNRNGEQRGGGTLCALPALWKAQPLVKNYQPC 347
 RESULT 4
 YKUS_CABEL
 ID_YKUS_CABEL STANDARD; PRT; 375 AA.


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; Z19153; CAD45582.1; -.
CC DR EMBL; Z27078; CAD45582.1; JOINED.
CC DR EMBL; Z27078; CAA81586.2; -.
CC DR EMBL; Z19153; CAA81586.2; JOINED.
CC DR EMBL; Z27078; CAB76843.1; -.
CC DR EMBL; AF220525; AAF43008.1; -.
CC DR PIR; S40996; S40996.
CC DR WormPep; K04H4.6a; CE25043.
CC DR WormPep; K04H4.6b; CE31813.
CC DR InterPro; IPR004947; DNase_II.
CC DR Pfam; PF03265; DNase_II; 1.
CC KW Hypothetical protein; Hydrolase; Signal; Alternative splicing.
CC FT SIGNAL 1 17 POTENTIAL.
CC FT CHAIN 18 378 HYPOTHELICAL PROTEIN K04H4.6.
CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT VARSPLIC 364 378 ALTVLLIIAIIITFFK -> VDALANNRYF (in isoform
CC a).
CC FT FTId-VSP 001304.
CC FT CONFLICT 40 40 L -> F (IN REF. 3).
CC FT CONFLICT 267 267 S -> P (IN REF. 3).
CC SQ SEQUENCE 378 AA; 43348 MW; E84D6C8F80EEF9 CRC64;

Query Match 24.0%; Score 462.5; DB 1; Length 378;
Best Local Similarity 31.3%; Pred.No.1.9e-32;
Matches 108; Conservative 62; Mismatches 158; Indels 17; Gaps 7;

QY 12 LLEFALSGVLGTPETISCRNEYGEAVDWFIYKLP---KRTSKASEAGLQYLYLDSTROT 68
DB 6 ILIVSLIGI-SNAAYCKDNNGSNVDWFVYKLPPLHWNHPDNPISNGTGFLYFDVNNKN 64
QY 69 WKSLYLINSTSALGRITLOHLYDTHNSTNDTAYLYNDGVPGSVNYSKQYGHAKGLLVW 128
DB 65 WKLMPOGMDVENNAVITYLQQYY--NSNNMTTFSYMYNDEWPDSTTWSNMSGHAKGVTFV 122
QY 129 NRTQGPWLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 198
DB 123 DQTGTFWLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 182
QY 189 CIPSTFWLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 248
DB 183 FNLPSQFANQFPVLSQKKNKYNKSPPLSTKVLKSLGQGHFHFHAKTGEWGDLYSDFV 242
QY 249 AQKLKTHLLAQTWQKK---KQELPSNCSLPYHYNYNKSIGVTSKSY---FSSQDHSKWC 302
DB 243 GPTLKSSIKVETWNHQSDEYNLPSVCD-PNHVQSTMSAKYIRLPYADYSSYEDHSKFV 301
QY 303 VSIKGSANR---WTCIGDLNLSLHQAALRGGGGFICTKNHVIYQAF 343
DB 302 VAYSESSAPPIPYVICIGDINRQSHQIHRRGGTMCYDQETYPQF 346

RESULT 7
ID DRN2 BOVIN STANDARD; PRT; 276 AA.
AC P56541; O46614;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)
DE (Lysosomal DNase II) (Fragment).
GN DNASE2 OR DNL2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;

```

```

QY 131 TQGFNLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 190
DB 119 TTGFNLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 178
QY 191 IPSTFWLHSPKPPVHVGYEPTSGRYGOTGICITFGYSQFEIDFOLLVLPNIYS 248
DB 179 IPEKYATREPTLNVEAKQSLSAQTQFWLSKPIKTVQVTLAYAKHKFKDGDINDLI 238
QY 249 AQKLKTHLLAQTW-QKKQELPSNCSLPYHYNYNKSIGVTSKSYFSSRQDHSKWCWSIKG 307
DB 239 SRQKVTFLAVESWLNGSGDDIHTCTSTSTQTHDVTETRTGLN-FASSKDHKSXWVS-NS 296
QY 308 SANRWICIGDLNLSLHQAALRGGGGFICTKNHVIYQAFKLYLRVGFCK 354
DB 297 QTNPIVCFGDMNRQSKLKRGGGALCIQNRNLWLYHSFVIQVEPCK 343

RESULT 6
ID YMV6 CAEEL STANDARD; PRT; 378 AA.
AC F34508; Q9NMQ0; Q9NCF8;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein K04H4.6 in chromosome III precursor.
GN K04H4.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Durbin R.; (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RX SEQUENCE OF 18-369 FROM N.A. (ISOFORM A).
RX MEDLINE=20363736; PubMed=10903447;
RA Krieser R.J., Eastman A.;
RT "Deoxyribonuclease II: Structure and chromosomal localization of the
RT murine gene, and comparison with the genomic structure of the human
RT and three C. elegans homologs."
RL Gene 252:155-162(2000).
CC -!- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=b;
CC IsoId=P34508-1; Sequence=Displayed;
CC Name=a;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the DNase II family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


RP VARIANTS SLOS MET-93; LEU-326; TRP-352 AND CYS-404.
RX MEDLINE-21109432; PubMed-11175299;
RA Witsch-Baumgartner M., Clara E., Löffler J., Menzel H.J., Seedorf U.,
RA Burn J., Gillesen-Kaesbach G., Hoffmann G.F., Fitzky B.U., Mundy H.,
RA Claydon P., Kelley R.I., Krzewska-Walasek M., Utermann G.,
RT "Frequency gradients of DHCR7 mutations in patients with Smith-Lemli-
RT Opitz syndrome in Europe: evidence for different origins of common
RT mutations.";
RL Eur. J. Hum. Genet. 9:45-50(2001).
RN [9].
RP VARIANTS SLOS LYS-448.
RX MEDLINE-22829205; PubMed-1294967;
RA Langius F.A., Waterham H.R., Romeijn G.J., Oostheim W., de Barse M.M.,
RA Dorland L., Duran M., Beemer F.A., Wanders R.J., Poll-The B.T.,
RT "Identification of three patients with a very mild form of
RT Smith-Lemli-Opitz syndrome.";
RL Am. J. Med. Genet. 122:24-29(2003).
CC -!- FUNCTION: Production of cholesterol by reduction of C7-C8 double
CC bond of 7-dehydrocholesterol (7-DHC).
CC -!- CATALYTIC ACTIVITY: Cholesterol + NADP(+) = cholesta-5,7-dien-3-
CC beta-ol + NADPH.
CC -!- PATHWAY: Cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -!- TISSUE SPECIFICITY: Most abundant in adrenal gland, liver, testis,
CC and brain.
CC -!- DISBAS: Defects in DHCR7 are the cause of Smith-Lemli-Opitz
CC syndrome (SLOS) [MIM:270400, 268670]; also known as SLO syndrome
CC or RSH syndrome. SLOS is an autosomal recessive frequent inborn
CC disorder of sterol metabolism with characteristic congenital
CC malformations and dysmorphias. All patients suffer from mental
CC retardation. Children with SLOS have elevated serum 7-
CC dehydrocholesterol (7-DHC) levels and low serum cholesterol
CC levels. SLOS occurs in relatively high frequency: approximately 1
CC in 20,000 to 30,000 births in populations of northern and central
CC European background. Historically, a clinical distinction often
CC was made between classic ("type I") SLOS and the more severely
CC affected ("type II") patients. There is, in reality, a clinical
CC and biochemical continuum from mild to severe SLOS.
CC -!- SIMILARITY: Belongs to the ERG4/ERG24 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF096305; AAC09766.1; -
DR EMBL; AF034544; AAC05086.1; -
EMBL; AF110060; AAD24762.1; -
DR EMBL; AF067127; AAD02816.1; -
DR EMBL; BC000054; AAH00054.1; -
DR EMBL; AF062481; AAC18345.1; -
DR Genbank; HGNC:2860; DHCR7.
DR MIM; 602858; -
DR MIM; 270400; -
DR MIM; 268670; -
DR InterPro; IPR001171; ERG4_ERG24.
DR Pfam; PF01222; ERG4_ERG24; 1.
DR PROSITE; PS01017; STEROL_REDUCT 1; 1.
DR PROSITE; PS01018; STEROL_REDUCT 2; 1.
KW Sterol biosynthesis; Cholesterol biosynthesis; Oxidoreductase; NADP;
Transmembrane; Endoplasmic reticulum; Disease mutation.
FT TRANSMEM 37 59 POTENTIAL.
FT TRANSMEM 151 173 POTENTIAL.
FT TRANSMEM 178 200 POTENTIAL.
FT TRANSMEM 264 286 POTENTIAL.
FT TRANSMEM 332 354 POTENTIAL.
FT TRANSMEM 414 436 POTENTIAL.
FT TRANSMEM 51 51 P -> S (in SLOS).
FT VARIANT /FTID=VAR_012717.

FT	VARIANT	93	93	T -> M (in SLOS).	
FT				/FTID=VAR_012718.	
FT	VARIANT	99	99	L -> P (in SLOS).	
FT				/FTID=VAR_012719.	
FT	VARIANT	119	119	H -> L (in SLOS).	
FT				/FTID=VAR_012720.	
FT	VARIANT	157	157	L -> P (in SLOS).	
FT				/FTID=VAR_012721.	
FT	VARIANT	244	244	G -> R (in SLOS).	
FT				/FTID=VAR_012722.	
FT	VARIANT	247	247	A -> V (in SLOS).	
FT				/FTID=VAR_012723.	
FT	VARIANT	248	248	W -> C (in SLOS).	
FT				/FTID=VAR_012724.	
FT	VARIANT	289	289	T -> I (in SLOS).	
FT				/FTID=VAR_012725.	
FT	VARIANT	326	326	V -> L (in SLOS).	
FT				/FTID=VAR_012726.	
FT	VARIANT	352	352	R -> W (in SLOS).	
FT				/FTID=VAR_012727.	
FT	VARIANT	380	380	C -> S (in SLOS).	
FT				/FTID=VAR_012728.	
FT	VARIANT	404	404	R -> C (in SLOS).	
FT				/FTID=VAR_012729.	
FT	VARIANT	410	410	G -> S (in SLOS).	
FT				/FTID=VAR_012730.	
FT	VARIANT	448	448	E -> K (in SLOS; mild).	
FT				/FTID=VAR_016975.	
FT	CONFLICT	5	5	S -> L (in REF. 2).	
FT	CONFLICT	14	14	S -> A (in REF. 5).	
FT	SEQUENCE	475 AA;	54489 MW;	7D726443834C4EEB CRC64;	
SQ					

Query Match 5.3%; Score 101.5; DB 1; Length 475;
Best Local Similarity 21.4%; Pred. No. 0.33;
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;

QY	6	LSVLSLFFA----	LSGLGTPPEISCRNEVEADWF-----	IFVKLPKRSKASE	53	
DB	40	LASVIFLLFAPFVYVFMACDQYSCALT-	GPVVDIVTGHARLSDIMAKTPTTRKAAQ	98		
QY	54	EAGL----	QYLYLDSRTQWTKSL-----	YLINSTRSALGRTLQH	89	
DB	99	LYTLWVTFVLLYTSLPDFCHFLSYVGGIOE	GAVTPAGVNVKQINGLOAWLLTHLW	158		
QY	90	LYDTNSTNDTAYLLINDGPGSVNYSROYG	HAKGLLVNRTQGFWLHSPKPPVHG	149		
DB	159	FANAHLSSWFSPTIIDNWIP-----	LLWCAN----	ILGYAVSTFAMVKGY	200	
QY	150	EYPTSGRRYGOTGICITFGYSQFEIDQLLV	QPNYSCE-----	IPSTFWKLYI	201	
DB	201	FFPTSARDCKFTG----	NFFYNTMMGIEP-----	NPRIGKWFDFKLFNRPQGVIAWTLIN	252	
QY	202	MPRMCANSSSLKIPVRYLAELHS--	AQGLNFVHFPAKSSFYTDIFTG	WIAQKLKTHLLAQ	259	
DB	253	L-----	SFAAK-----	QRELHSHVNAWLVNVLQ-AIYVIDFF--	W-----NE	288
QY	260	TWOKKQEL-----	PSNCSLP-----	Y	276	
DB	289	TWYLTIDICHDFGWLWGDVWLPYLYTIQGL	YLVVHPVQLSTPHAVGVLLGLVGY	348		
QY	277	HYVNI-----	KSIGVTSKYESS--	RQDHSKWCYS-IKGSANRWTC	314	
DB	349	YIFRVANQKDLFRRTDRCGLWGKPKVICS	YTSADGQRHSHKLLVSGFWGARHFN	Y	408	
QY	315	IGDLNRSLLHQALR-GGGFICTKNHYIQAF-	-----	HKLYLRYG	351	
DB	409	VGLDMSLAYCLACGGHLLPFYIYVAILL	THRECLREHRCASKYG	456		

RESULT 9
CSAA_BACUD
ID CSAA_BACUD
AC Q45760; STANDARD; PRT; 1385 AA.

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Pesticidal crystal protein cry5Aa (Insecticidal delta-endotoxin
 DE Cry5Aa) (Crystalline entomocidal protoxin) (152 kDa crystal protein).
 GN CRY5AA OR CRYVA(A) OR CRVVA.
 OS Bacillus thuringiensis (subsp. darmstadtensis).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=132264;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-18243 / PS17;
 RA Slack A.J., Schwab G.E., Payne J.M.;
 RT "Genes encoding nematode-active toxins cloned from Bacillus
 RT thuringiensis isolate PS17";
 RL Patent number US5281530, 25-JAN-1994.
 CC -1- FUNCTION: ENDOTOXIN WITH NEMATOCIDAL ACTIVITY.
 CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -1- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L07025; AAA67694.1; --
 DR PIR; T18213; T18213.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR Pfam; PF03944; endotoxin_C_1.
 DR Pfam; PF03945; endotoxin_N_1.
 KW Toxin; Sporulation.
 SQ SEQUENCE 1385 AA; E29FF1FC799DE95 CRC64;
 Query Match 4.9%; Score 94; DB 1; Length 1385;
 Best Local Similarity 20.58; Pred. No. 5.6;
 Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;
 Qy 31 EYGEADVWTFYKLPKRTSKASEBAGLYLYLSTR-----QTWKN----- 71
 Db EVQDSVETRLYGQLP-----AVDQAGPNYVSIDSSNPLIQINMDTKTPPQAGSGMNTNL 454
 Qy 72 -----SLYLINSTRSALGR-----TLQHLVDYTHNSTNDTAYLIYNDGVP--GSVNY 115
 Db MRGSVSGLSFLQDGTHTLSAGMGGFADTYSLSLPAH-----YLSLYGTPYQTSNDY 507
 Qy 116 SROYGHAKGLVWNRRTQGFMLIHSPVK---PPPVHGYEYPTSGRRYQGTGTCITFGYSQF 172
 Db SCHVGALGVV-----STPQEAFLPIIG-----QPDEQGNVSTMGF-PF 545
 Qy 173 ERIDFOLLVLQPNYSCFIPSTFHWKLIYMPRCMANSSSLKIPVRYLAELHS----- 224
 Db EKASYGTVVKEWMLNGANAKMLSPQSGTIGIPITNVTSGEYQIRCRYASNDNTNFFNVDT 605
 Qy 225 -----AQLNFVHFHAKSSFYTDDIFTGWIQAOKLKTLLAQ---- 259
 Db GGANPIFQINFASVTVDNNTGVQGVVYVKSATDTSNSTEIPAKTINVHLTNQSSD 665
 Qy 260 TWQKKQBLPNCISLPYH-VYNIKSIGVTSKSYFSSRQDHSKWCVSINKGSANRWTCIGDL 318
 Db VFLDRIFIPSLPLIYHGSYNTSS-GADVDLWSSNNY--YDLIVNGQANS-----SSI 718
 Qy 319 NRSLLHQAURG 328
 Db ASSMHLNKG 728

RESULT 10
 RRPL_HRSVA STANDARD; PRT; 2165 AA.
 ID RRPL_HRSVA
 AC P28887;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
 DE (L protein).
 GN L.
 OS Human respiratory syncytial virus (strain A2).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11259;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91272488; PubMed=2053282;
 RA Stec D.S., Hill M.G. III, Collins P.L.;
 RT "Sequence analysis of the polymerase L gene of human respiratory
 RT syncytial virus and predicted phylogeny of nonsegmented
 RT negative-strand viruses";
 RL Virology 183:273-287(1991).
 CC -1- FUNCTION: Probable component of the active polymerase. It may
 CC function in mRNA synthesis, capping, methylation and poly(A)
 CC synthesis of newly synthesized viral mRNAs, RNA editing of the P
 CC gene transcript, and protein kinase activity.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -1- SIMILARITY: Belongs to the paramyxoviruses L protein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M75730; AAA47418.1; --
 DR PIR; A40317; RRNA2A.
 DR InterPro; IPR007098; RNA_pol_monon.
 DR InterPro; IPR001016; Viral_RNA_pol_L.
 DR Pfam; PF00946; Paramyx_RNA_pol_1.
 KW Transferase; RNA-directed RNA polymerase.
 SQ SEQUENCE 2165 AA; 250384 MW; 5A3DC8A0805305197 CRC64;
 Query Match 4.9%; Score 94; DB 1; Length 2165;
 Best Local Similarity 21.5%; Pred. No. 9.7;
 Matches 58; Conservative 34; Mismatches 92; Indels 86; Gaps 12;
 Qy 112 SSVNSROYGHAKGLVWNRRTQGFMLIHSPVKFPVPHGYEYPTS-----GRYQGTG-- 163
 Db 1333 SVNYLHR-----LTVSSRPCEF--PASIPAYRTTN-YHFDTSINRLITKYGDEDLDI 1383
 Qy 164 ---CITFGYSQFBEIDFOLLVLQPNYSCFIPSTFHWKLIYMPRCMANSSSLKIPVRYL 219
 Db 1384 VFQNCISFGLSLMSVVB-QFTNVCPN-----RIILIFK----- 1415
 Qy 220 AELHSAQGLNFVHFAKSSFYTDDIFTGWIQAOKLKTLLAQWQKKQBLPNCISLPYHY 279
 Db 1416 -----LNEIHLMKPPIFTGDV-----DIHKLKQVIRQHMFLPKISLTQYV- 1457
 Qy 280 NIKSIGVTSKSYFSSRQDHSKWCVSINKGS-----ANRWTCIGDL---NRSLLH 323
 Db 1458 ---ELFJUNTKRSGSHVNSNLLIAHKLSIDYFHTNYLTSTNLACHWILLIQLMKDSKGI 1514
 Qy 324 QALRGGGFICTKKNHYIYQAFLKLYRYCFC 353
 Db 1515 EKDNGEGYITDHFMLNKVFNAYKTYLLC 1544

Db	32	PEWSGLHQHGD	-----YLAQRLAIVDPASGDQPLFNEPNISVITVNGEIYNHEDLRK	85
Qy	82	ALGR	-----TLQHLIYDTHN	
Db	86	QUSNHTFRGSDCVIAHLYEEYGBDFVMDLGDIFSVPLDTRDYSIIVARDAI	-----STNDTAYLIYNDGVPGSVNY	115
Qy	116	SRQYGHAKGLIWNRTQGFMLHSPK	-----PPVHGVEYFPTSG-RRYGOTGIC	164
Db	140	---GVTSIYICWGLDGSVWISSEKGLNDDCEHFECFFPHGLHYSKSGSFFRW	-----	189
Qy	165	ITFGYSQFEIDFQLLVLQPNIYSFIFSTFH	-----WKLIIYMPRCAN	208
Db	190	-----YNPSYSEAI	-----	232
Qy	209	---SSSL-KIPVRVYLAELHSAQ	-----GLNFVHFAPKSFYTDIDFTG-WIAQKLKT--HLIA	258
Db	233	GGLDSSLVASITSRYLATTKAAEQWGSKLHSCVGLGEGSPDLKAGKEVADYLGTVHHEFT	-----	292
Qy	259	QTWOKKQKELPNSGLPYHV	-----YNIKISGIVTSKSYFSSRQHS	313
Db	293	FTVQDGDIAIB	-----DVIVHVTYDVTSTRAPFMFLMRKIKSLGKVVWISGEGS	344
Qy	314	CIGDLNRLSHQALRGGGFICTKNHYIYQAFHK	345	
Db	345	-----DEIFGG	-----YLY--FHK	356
RESULT 12				
SCAP_CRIGR STANDARD; PRT; 1276 AA.				
ID	SCAP_CRIGR	STANDARD;	PRT;	1276 AA.
AC	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Sterol regulatory element binding protein cleavage-activating protein			
DE	(SREBP cleavage-activating protein) (SCAP).			
GN	Cricetulus griseus (Chinese hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Cricetulus.			
OX	NCBI_Taxid:10029;			
FN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-443.			
RC	TISSUE=Ovary;			
RC	MEDLINE=97053781; PubMed=8898195;			
RA	Hua X., Nohruff A., Goldstein J.L., Brown M.S.;			
ET	"Sterol resistance in CHO cells traced to point mutation in SREBP			
RT	cleavage-activating protein.";			
RL	Cell 87:415-426(1996).			
RL	[2]			
RP	POST-TRANSLATIONAL MODIFICATIONS, AND MUTAGENESIS OF TYR-298.			
RC	TISSUE=Ovary;			
RC	MEDLINE=89432216; PubMed=10500160;			
RA	Nohruff A., DeBose-Boyd R.A., Sheek S., Goldstein J.L., Brown M.S.;			
RT	"Sterols regulate cycling of SREBP cleavage-activating protein (SCAP)			
RL	between endoplasmic reticulum and Golgi.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:11235-11240(1999).			
RL	[3]			
RP	TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES ASN-263; ASN-590 AND ASN-641.			
RC	MEDLINE=98307971; PubMed=9642295;			
RA	Nohruff A., Brown M.S., Goldstein J.L.;			
RT	"Topology of SREBP cleavage-activating protein, a polytopic membrane			
RL	protein with a sterol-sensing domain.";			
RL	J. Biol. Chem. 273:17243-17250(1998).			
RL	[4]			
RP	SUBCELLULAR LOCATION.			
RC	MEDLINE=20449077; PubMed=10896675;			
RA	Yang T., Goldstein J.L., Brown M.S.;			
RT	"Overexpression of membrane domain of SCAP prevents sterols from			
RL	inhibiting SCAP. SREBP exit from endoplasmic reticulum.";			
RL	J. Biol. Chem. 273:29881-29886(2000).			
CC	-I- FUNCTION: Sterol sensor. Necessary for the proteolytic activation			

of SREBPs by site-1 protease in the Golgi.
 CC -!- SUBUNIT: In a tight complex with SREBs.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Moves from the
 CC endoplasmic reticulum to the Golgi in the absence of sterols.
 CC -!- SIMILARITY: Contains 1 sterol sensor (SSD) domain.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U67060; AAB19103.1; -.
 CC PIR: T18526.
 CC GO: GO:0005783; C:Endoplasmic reticulum; ISS.
 CC GO: GO:0005794; C:Golgi apparatus; ISS.
 CC GO: GO:0016021; C:Integral to membrane; ISS.
 CC GO: GO:0003754; F:Chaperone activity; NAS.
 CC GO: GO:0045541; P:Positive regulation of cholesterol biosynth. . . ; ISS.
 CC GO: GO:0045716; P:Positive regulation of low-density lipoprot. . . ; ISS.
 CC GO: GO:0006994; P:sterol depletion response, SREBP target gen. . . ; ISS.
 CC InterPro: IPR000731; SSD 5TM.
 CC InterPro: IPR001680; WD40.
 CC Pfam: PF00400; WD40; 4.
 CC SMART: SM00320; WD40; 6.
 CC PROSITE: PS0156; WD REPEATS 1; 1.
 CC PROSITE: PS00678; WD REPEATS 2; 1.
 CC PROSITE: PS50082; WD REPEATS 2; 1.
 CC PROSITE: PS50294; WD REPEATS REGION; 1.
 CC Lipid metabolism; Cholesterol metabolism; Transmembrane;
 CC Endoplasmic reticulum; Golgi stack; Repeat; WD repeat; Glycoprotein.
 CC
 CC DOMAIN 1 18
 CC TRANSFEM 19 39
 CC DOMAIN 40 279
 CC TRANSFEM 280 300
 CC DOMAIN 301 312
 CC TRANSFEM 313 333
 CC DOMAIN 334 344
 CC TRANSFEM 345 365
 CC DOMAIN 366 401
 CC TRANSFEM 402 422
 CC DOMAIN 423 423
 CC TRANSFEM 424 444
 CC DOMAIN 445 518
 CC TRANSFEM 519 539
 CC DOMAIN 540 708
 CC TRANSFEM 709 729
 CC DOMAIN 730 1276
 CC TRANSFEM 284 442
 CC REPEAT 771 811
 CC REPEAT 771 811
 CC REPEAT 949 999
 CC REPEAT 1002 1039
 CC REPEAT 1074 1111
 CC REPEAT 1114 1152
 CC REPEAT 1155 1192
 CC REPEAT 1194 1232
 CC POLY-LEU.
 CC POLY-ARG.
 CC N-LINKED (GLCNAC. . .)
 CC N-LINKED (GLCNAC. . .)
 CC N-LINKED (GLCNAC. . .)
 CC Y->C: ABOLISHES STEROL-RESPONSE.
 CC D->N: ABOLISHES STEROL-RESPONSE.
 CC MUTAGEN 443 443
 CC SEQUENCE 1276 AA; 139512 MW; A8693F7157FF5FEC CRC64;

Query Match 4.7%; Score 90; DB 1; Length 1276;
 Best Local Similarity 17.9%; Pred. No. 11;
 Matches 56; Conservative 50; Mismatches 102; Indels 104; Gaps 13;

29 RNEYCEAVDWFYFKLPKRTSKASEAGLYLDSTRQTNWNSLYLINSTRSALGRTIQ 88

Db 75 QCEPSQPEWYV-----CAPVAYTQQIFVKSSVSPWHKNLLAVDVRLPLSRAFQ 124
 QY 89 HLYDTHN-----STNDTAVLIYNDGVPGSVNVSYRGYGHAKGLLVNRTQGFWLHS 139
 Db 125 LVEEIRNHLVRSSGTYKSEEVCLQVTDLLPG-----LRK 159
 QY 140 VPKFPVGHVEYPTSGRRYGTGTCITFGYSQFEEIDFQLLVLPQNIYSFISTFHWKL 199
 Db 160 LRNLLEHG-----CLLSPGNQWDERFHADPDI-----GTIHQ-- 197
 QY 200 IYMPRCANSSSLK-----IPVRYLAELSAQGLNFVHFAKSFYTDIDFTGWLAKLKT 254
 Db 198 -HEPKTLQTSATLKDLFGVPGKY-----SGVSLYTRKRTVSYTTL-----VFQYHA 245
 QY 255 HLLAQTWQKKQQLPS-NCSLP-----YHVYNIKSIGVT-----SKSYFSRRQ 296
 Db 246 KFLSSLRARMLLHPSPNCRAENLVHVFKEIGIAELPLVTTVILLPAYIYFSTRK 305
 QY 297 ---DHGKWCVSI 305
 Db 306 IDAVKSKWGLAL 317
 RESULT 13
 PTPG HUMAN
 ID PTPG HUMAN STANDARD; PRT; 1445 AA.
 AC P23470; Q15623;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Protein-tyrosine phosphatase gamma precursor (EC 3.1.3.48) (R-PTP-
 DE gamma).
 DE PTPRG.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93180796; PubMed=8382771;
 RA Barnea G., Silvenoinen O., Shaanan B., Honegger A.M., Canoll P.D.,
 RA D'Ustachio P., Morse B., Levy J.B., Laforgia S., Huebner K.,
 RA Musacchio J.M., Sap J., Schleisinger J.;
 RT "Identification of a carbonic anhydrase-like domain in the
 RT extracellular region of RPTP gamma defines a new subfamily of
 RT receptor tyrosine phosphatases.";
 RL Mol. Cell. Biol. 13:1497-1506(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96429999; PubMed=8833149;
 RA Kastury K., Ohta M., Lasota J., Moir D., Dorman T., Laforgia S.,
 RA Druck T., Huebner K.;
 RT "Structure of the human receptor tyrosine phosphatase gamma gene
 RT (PTPRG) and relation to the familial RCC t(3;8) chromosome
 RT translocation.";
 RL Genomics 32:225-235(1996).
 RN [3]
 RP SEQUENCE OF 836-1445 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91006018; PubMed=2170109;
 RA Krueger N.X., Streuli M., Saito H.;
 RT "Structural diversity and evolution of human receptor-like protein
 RT tyrosine phosphatases.";
 RL EMBO J. 9:3241-3252(1990).
 RN [4]
 RP SEQUENCE OF 874-1118 AND 1175-1409 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=90384936; PubMed=2169617;
 RA Kaplan R., Morse B., Huebner K., Croce C., Hawk R., Ravera M.,
 RA Ricca G., Jaye M., Schleisinger J.;
 RT "Cloning of three human tyrosine phosphatases reveals a multigene
 RT family of receptor-linked protein-tyrosine-phosphatases expressed in

RA Dybvig K.;
 RT "DNA polymerase III of Mycoplasma pulmonis: isolation and
 RL characterization of the enzyme and its structural gene, polC.";
 RL Mol. Microbiol. 13:843-854(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UBA CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
 RA Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen
 RT Mycoplasma pulmonis";
 RL Nucleic Acids Res. 29:2145-2153(2001).
 CC -!- FUNCTION: Required for replicative DNA synthesis. This DNA
 CC polymerase also exhibits 3' to 5' exonuclease activity.
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA](N).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-C family. PolC
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 DR EMBL; U06933; AAA18178.1; --
 DR EMBL; AL445565; CAC13848.1; --
 DR PIR; C90596; C90596.
 DR PIR; S54697; S54697.
 DR MYPULIST; MYPUL 6750; --
 DR HAMAP; MF 00356; --
 DR InterPro; IPR006054; DnaQ.
 DR InterPro; IPR006055; Exonuclease.
 DR InterPro; IPR003141; PIP N.
 DR InterPro; IPR006308; PolC_gram_pos.
 DR Pfam; PF00929; Exonuclease; 1.
 DR Pfam; PF02231; PIP N; 1.
 DR SMART; SM00479; EXOIII; 1.
 DR SMART; SM00481; POLIIIAC; 1.
 DR TIGRFAMS; TIGR00573; dnaq; 1.
 DR TIGRFAMS; TIGR01405; polC_Gram_pos; 1.
 DR Transferrase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
 KW Nuclease; Exonuclease; Complete proteome.
 FT DOMAIN 404 570 EXONUCLEASE.
 FT CONFLICT 786 787 NP -> S (IN REF. 1).
 FT CONFLICT 1233 1233 R -> KE (IN REF. 1).
 SQ SEQUENCE 1435 AA; 166374 MW; 9E6527A88725503F CRC64;
 Query Match 4.6%; Score 88.5; DB 1; Length 1435;
 Best Local Similarity 22.8%; Pred. No. 17;
 Matches 75; Conservative 42; Mismatches 121; Indels 91; Gaps 22;
 QY 18 SGVLGTPRIS---CRNEYGEAVDMFVFKLPKRTSKASEAGLQVLYLDSTQVW-NKSL 73
 Db 1134 TGAIGPEFGTGFVRQMLSEA-----KPKTFADLVISIGLSH-----GTDVWINNAH 1180
 QY 74 YLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYS--ROYGHAKGLLVWNRK 131
 Db 1181 YIIQS-----LQKTLQVLI---SCRDDIMVDLKKGVFDLSFTIMEQVKGKLSLEQKR 1233
 QY 132 Q-----GF--WLTHSVPK-----PPVHGVEYPTSGR-----YGTGICITFGYSQFEEI 175
 Db 1234 KLEHGIENWYTESMEKIKYMPKKAHATAYVLMAWVAFYKVVYPLEYVATYFSTRTEFF 1293
 QY 176 DFQL-----LVLPNIYSCFISTFWKLIYMPRCANSSSL-----KIPVRYLAELHS 224
 Db 1294 DIEIMSKDKLTLESKIK-----ELAYRENIL-RNDNQLTKTKNTLPTLYIANEMK 1342

QY 225 AQLNFVHF-AKSFYTDIFTGWIQAQKLKTHLLAQTWQKKQELPSNCSLPHVHNKIS 283
 Db 1343 ARGFNQINIMKISLAND-----WIIDK-----NSKSLIP-----PENV--IDG 1379
 QY 284 IGVT-SKSYFSSRQDHSKWCVSIGKSANR 311
 Db 1380 LGETLQAQIVDSRNE--KEFLSVDFINR 1406
 RESULT 15
 POLG DEN2P STANDARD; PRT; 3388 AA.
 AC P12823; Q88646; Q88647; Q88648; Q88649; Q88650; Q88651; Q88652;
 AC Q88653; Q88654; Q88655;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope glycoprotein M); Major envelope protein E;
 DE Nonstructural proteins NS1, NS2A, NS2B, NS4A and NS4B;
 DE Protease/helicase (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase
 DE (EC 2.7.7.48) (NS5)].
 OS Dengue virus type 2 (strain PR159/S1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 CX NCBI_TaxID=11066;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88101365; PubMed=2827375;
 RA Hahn Y.S., Galler R., Hunkapiller T., Dalrymple J.M., Strauss J.H.,
 RA Strauss E.G.;
 RT "Nucleotide sequence of dengue 2 RNA and comparison of the encoded
 RT proteins with those of other flaviviruses.";
 RL Virology 162:167-180(1988).
 RN [2]
 RP C-TERMINUS OF M; E AND NS1.
 RC STRAIN=New-Guinea;
 RX MEDLINE=89299482; PubMed=2741348;
 RA Wright P.J., Cauchi M.R., Ng M.L.;
 RT "Definition of the carboxy termini of the three glycoproteins
 RT specified by dengue virus type 2.";
 RL Virology 171:61-67(1989).
 CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
 CC hydrophobic, suggesting a possible membrane-related function.
 CC NS3 and NS5 may play a role in the viral RNA replication.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1',
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
 CC lipoprotein envelope. The envelope consists of two proteins:
 CC protein M and glycoprotein E. The nucleocapsid is a complex of
 CC protein C and mRNA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 DR EMBL; M19197; AAA42962.1; --
 DR MEROPS; S07_002; --
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001122; Flavi_capsidC.
 DR InterPro; IPR000336; Flavi_Glycoprote.
 DR InterPro; IPR000069; Flavi_M.
 DR InterPro; IPR001157; Flavi_NS1.
 DR InterPro; IPR000752; Flavi_NS2A.
 DR InterPro; IPR000487; Flavi_NS2B.

DR	InterPro: IPR000404; Flavi_NS4A.	
DR	InterPro: IPR001528; Flavi_NS4B.	
DR	InterPro: IPR000208; Flavi_NS5.	
DR	InterPro: IPR002535; Flavi_proprep.	
DR	InterPro: IPR007110; Ig-like.	
DR	InterPro: IPR001850; Peptidase_S7.	
DR	InterPro: IPR007095; RNA_pol_DS_PS.	
DR	InterPro: IPR007094; RNA_pol_PSVir.	
DR	InterPro: IPR002877; RrmJ_FtsJ.	
DR	Pfam: PF01003; Flavi_capsid; 1.	
DR	Pfam: PF02832; Flavi_glycop_C; 1.	
DR	Pfam: PF00869; Flavi_glycoprot; 1.	
DR	Pfam: PF00949; Flavi_helicase; 1.	
DR	Pfam: PF01004; Flavi_M; 1.	
DR	Pfam: PF00948; Flavi_NS1; 1.	
DR	Pfam: PF01005; Flavi_NS2A; 1.	
DR	Pfam: PF01002; Flavi_NS2B; 1.	
DR	Pfam: PF01350; Flavi_NS4A; 1.	
DR	Pfam: PF01349; Flavi_NS4B; 1.	
DR	Pfam: PF00972; Flavi_NS5; 1.	
DR	Pfam: PF01570; Flavi_proprep; 1.	
DR	Pfam: PF01728; FtsJ; 1.	
DR	ProDom: PD003556; Flavi_glycoprote; 1.	
DR	ProDom: PD001496; Flavi_NS1; 1.	
DR	ProSITE; PS00690; DEAH_ATP_HELICASE; FALSE NEG.	
DR	Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;	
KW	Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;	
KW	ATP-binding; Transmembrane; Nonstructural protein.	
FT	CHAIN 1 114	
FT	PROPEP 115 205	
FT	CHAIN 206 280	ENVELOPE GLYCOPROTEIN M.
FT	CHAIN 281 775	MAJOR ENVELOPE PROTEIN F.
FT	CHAIN 776 1127	NONSTRUCTURAL PROTEIN NS1.
FT	CHAIN 1128 1345	NONSTRUCTURAL PROTEIN NS2A.
FT	CHAIN 1346 1475	NONSTRUCTURAL PROTEIN NS2B.
FT	CHAIN 1476 2090	PROTEASE/HELICASE (NS3).
FT	CHAIN 2091 2376	NONSTRUCTURAL PROTEIN NS4A.
FT	CHAIN 2377 2488	NONSTRUCTURAL PROTEIN NS4B.
FT	CHAIN 2489 3388	RNA-DIRECTED RNA POLYMERASE (NS5).
FT	NP_BIND 1668 1675	ATP (POTENTIAL).
FT	SITE 1759 1762	DEAH BOX.
FT	TRANSMEM 43 59	POTENTIAL.
FT	TRANSMEM 101 117	POTENTIAL.
FT	TRANSMEM 268 284	POTENTIAL.
FT	TRANSMEM 727 743	POTENTIAL.
FT	TRANSMEM 757 773	POTENTIAL.
FT	DISULFID 283 310	BY SIMILARITY.
FT	DISULFID 340 396	BY SIMILARITY.
FT	DISULFID 354 385	BY SIMILARITY.
FT	DISULFID 372 401	BY SIMILARITY.
FT	DISULFID 465 565	BY SIMILARITY.
FT	DISULFID 582 613	BY SIMILARITY.
FT	CARBOHYD 183 183	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 347 347	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 433 433	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 905 905	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 982 982	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1134 1134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1174 1174	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2298 2298	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2302 2302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2454 2454	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2482 2482	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 3388 AA; 379213 MW; 79B7C87B364D2DF CFC64.	

QY	253	KTHLLAQTWOKKQOEI	PSNCSLPYHYVNIKS	IGVTSKSYFSSRQDSK	SCVCSIKGSANRW	312
		:	:	:	:	:
Db	3303	KWILENPNMDEKTPV	ESWEEI	-----	YLGKED--QWCGSLIGTSRA	3345
		:	:	:	:	:
QY	313	TCIGDLNRSIHQ	324			
		:	:	:	:	:
Db	3346	TWAKNIQTAINQ	3357			

Search completed: October 14, 2004, 00:29:05
Job time : 26.3966 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 23:51:40 ; Search time 56.2616 Seconds
(without alignments)
1985.252 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLRLRYGFCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1927	100.0	354	11 Q9QY48	Q9QY48 mus musculus
2	1924	99.8	354	11 Q8C589	Q8C589 mus musculus
3	1373	71.3	356	11 Q9QZK9	Q9QZK9 rattus norv
4	1293.5	67.1	357	4 Q9NQW3	Q9NQW3 homo sapien
5	1293.5	67.1	361	4 Q8WZ79	Q8WZ79 homo sapien
6	692.5	35.9	375	12 Q9Q759	Q9Q759 fowlpox vir
7	567	29.4	153	4 Q8WZ80	Q8WZ80 homo sapien
8	566	29.4	350	11 Q9QZK8	Q9QZK8 rattus norv
9	455	23.6	232	12 Q9J5H0	Q9J5H0 fowlpox vir
10	444	23.0	366	5 Q9VED8	Q9VED8 drosophila
11	169.5	8.8	338	5 Q9BJL8	Q9BJL8 trichinella
12	151.5	7.9	339	5 Q8GQP5	Q8GQP5 trichinella
13	148.5	7.7	342	5 Q8GQP4	Q8GQP4 trichinella
14	111.5	5.8	832	10 Q7XQ00	Q7XQ00 oryza sativ
15	106	5.5	622	2 Q45155	Q45155 bacteroides
16	106	5.5	1115	12 Q9QMH2	Q9QMH2 parvo-like

17	104	5.4	276	5	Q962V2	Q962V2 trichinella
18	104	5.4	344	5	Q27073	Q27073 trichinella
19	101.5	5.3	315	5	Q27076	Q27076 trichinella
20	99	5.1	735	16	Q8XNF9	Q8XNF9 clostridium
21	97.5	5.1	1886	2	Q7X504	Q7X504 leptospira
22	97.5	5.1	2499	2	Q8GFA0	Q8GFA0 photorhabdu
23	97	5.0	344	5	Q8MTB3	Q8MTB3 trichinella
24	97	5.0	419	10	Q9SMI9	Q9SMI9 cuphea lanc
25	97	5.0	646	5	Q8ICY6	Q8ICY6 plasmodium
26	95	4.9	508	16	Q81ZA6	Q81ZA6 bacillus an
27	95	4.9	1770	5	Q813A3	Q813A3 plasmodium
28	94	4.9	172	2	Q9L910	Q9L910 escherichia
29	94	4.9	2165	12	O99721	O99721 respiratory
30	94	4.9	2165	12	O41355	O41355 human respi
31	94	4.9	2165	12	Q91W88	Q91W88 human respi
32	94	4.9	2165	12	Q82027	Q82027 human respi
33	94	4.9	2165	12	O41356	O41356 human respi
34	94	4.9	2165	12	Q82021	Q82021 human respi
35	94	4.9	2165	12	P90197	P90197 human respi
36	93.5	4.9	712	12	Q9DUC1	Q9DUC1 tt virus. o
37	93.5	4.9	716	17	Q9HLF9	Q9HLF9 thermoplasm
38	93	4.8	756	16	Q8DKB1	Q8DKB1 synechococc
39	92.5	4.8	320	12	Q9Q14	Q9Q14 shope fibro
40	91.5	4.7	399	17	Q8ZVF1	Q8ZVF1 pyrobaculum
41	91.5	4.7	774	16	Q92CV7	Q92CV7 listeria in
42	91	4.7	763	12	Q99A78	Q99A78 tt virus. o
43	90.5	4.7	406	10	Q94147	Q94147 oryza sativ
44	90.5	4.7	518	2	O87771	O87771 legionella
45	90	4.7	504	2	Q9ZTP5	Q9ZTP5 rhizobium 1

ALIGNMENTS

RESULT 1

ID	Q9QY48	PRELIMINARY;	PRT;	354 AA.
AC	Q9QY48;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Deoxyribonuclease DLAD.			
GN	DNASE2B OR DLAD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6; TISSUE=Liver;			
RX	MEDLINE=99428663; PubMed=10497274;			
RA	Shiokawa D., Tanuma S.;			
RT	"DLAD, a novel mammalian divalent cation-independent endonuclease with			
RT	homology to DNase II.";			
RL	Nucleic Acids Res. 27:4083-4089 (1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RX	MEDLINE=21556924; PubMed=11700027;			
RA	Shiokawa D., Tanuma S.;			
RT	"Isolation and Characterization of the DLAD/Dlad Genes, Which Lie			
RT	Head-to-Head with the Genes for Urate Oxidase.";			
RL	Biochem. Biophys. Res. Commun. 288:1119-1128 (2001).			
DR	EMBL; AF128888; AAF05082.1; -.			
DR	EMBL; AF334608; AAL34450.1; -.			
DR	EMBL; AF334603; AAL34450.1; JOINED.			
DR	EMBL; AF334604; AAL34450.1; JOINED.			
DR	EMBL; AF334605; AAL34450.1; JOINED.			
DR	EMBL; AF334606; AAL34450.1; JOINED.			
DR	EMBL; AF334607; AAL34450.1; JOINED.			
DR	MGD; MGI:1913283; Dnae2b.			
DR	GO; GO:0005737; C:cytoplasm; IDA.			
DR	GO; GO:0005576; C:extracellular; IDA.			

Fri Oct 15 10:01:04 2004

```
DR GO:0004520; F:endonucleolytic activity; IDA.
DR InterPro: IPR004947; DNase_II.
DR Pfam: PF03265; DNase_II; 1.
SQ SEQUENCE 354 AA; 40793 MW; 40A5DD6E9AD278A1 CRC64;

Query Match 100.0%; Score 1927; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 4e-167;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60

QY 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120
Db 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120

QY 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180
Db 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180

QY 181 VLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240
Db 181 VLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240

QY 241 DDIFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300
Db 241 DDIFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300

QY 301 WCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 354
Db 301 WCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 354

RESULT 2
Q8C589 PRELIMINARY; PRT; 354 AA.
AC Q8C589;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Deoxyribonuclease II beta.
GN DNASE2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK079224; BAC37579.1; -.
DR MGD; MGI:1913283; DNase2b.
DR GO:0005737; C:cyttoplasm; IDA.
DR GO:0005576; C:extracellular; IDA.
DR GO:0004520; F:endonucleolytic activity; IDA.
DR InterPro: IPR004947; DNase_II.
DR Pfam: PF03265; DNase_II; 1.
SQ SEQUENCE 354 AA; 40765 MW; 40AFB179AD5D428 CRC64;

Query Match 99.8%; Score 1924; DB 11; Length 354;
Best Local Similarity 99.7%; Pred. No. 7.5e-167;
Matches 353; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60

QY 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120
Db 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120

QY 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180
Db 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180

QY 181 VLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240
Db 181 VLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240

QY 241 DDIFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300
Db 241 DDIFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300

QY 301 WCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 354
Db 301 WCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 354

RESULT 3
Q9QZK9 PRELIMINARY; PRT; 356 AA.
AC Q9QZK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=20025354; PubMed=10558878;
RA Tanuma Si.; Shiohawa D.;
RT "Cloning of a cDNA encoding a rat DNase II-like acid DNase.";
RL Biochem. Biophys. Res. Commun. 265:395-399 (1999).
DR EMBL; AF178974; AAF13596.1; -.
DR PIR; JC7131; JC7131.
DR GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO:0006259; P:DNA metabolism; IEA.
DR InterPro: IPR004947; DNase II.
DR Pfam: PF03265; DNase II; 1.
SQ SEQUENCE 356 AA; 40472 MW; 5CEB359F42DD6741 CRC64;

Query Match 71.3%; Score 1373; DB 11; Length 356;
Best Local Similarity 71.0%; Pred. No. 1.2e-116;
Matches 252; Conservative 34; Mismatches 67; Indels 2; Gaps 1;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60

QY 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120
Db 61 YLDSTRTQWNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNSROYG 120

QY 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180
Db 121 HAKGLLVNWRTOGFWLIHSPKPPVHGYPYPTSGRRYGTGICITFGYQFSEIDPQLL 180

QY 179 LIVLOPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSF 238
Db 179 LIVLOPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSF 238

QY 239 YTDIDFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 298
Db 239 YTDIDFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 298

QY 241 YTDIDFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300
Db 241 YTDIDFTGMIAGKLTLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKYSFSSRDHSHK 300

QY 299 SKWCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 353
Db 299 SKWCVSIKGSANRWTICIGDLNLSHQAALRGGFICTKNHYIYQAFHKLILRYGFCK 353
```



```

Db 301 SKWCYSTKDSQARWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 355
|||||
RESULT 4
Q9NQW3 PRELIMINARY; PRT; 357 AA.
AC Q9NQW3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II beta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21272509; PubMed=11376952;
RA Krieser R.J., Maclellan K.S., Park J.P., Eastman A.;
RT "The cloning, genomic structure, localization, and expression of human
RL deoxyribonuclease II beta.";
RL Gene 269:205-216(2001).
DR EMBL; AF274571; AA76693.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 357 AA; 41197 MW; 5565D2036BD20955 CRC64;

Query Match 67.1%; Score 1293.5; DB 4; Length 357;
Best Local Similarity 66.4%; Pred. No. 2.1e-109;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLISLFFALSGVLGTPETISCRNEYGEAVDWFYKLPKRTSKASBEAGLQYL 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MWARLLRTSFALLFLGLFGLGATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 YLDSTRQTWNKSLYLINSTRSALGRITLQHLIDYTHNS--TNDTAYLIYNDGVPKPNYSRKY 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YLDSTRSWEKSEQLMNDYKSLVGRITLQQLYEAAYASKNNNTAYLIYNDGVPKPNYSRKY 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 GHAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 GHAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 GHTKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 GHTKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAOGLNVHFAKSS 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 178 QLLVLPQNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAOGLNVHFAKSS 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 FYTDDIFTGWIAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 FYTDDIFTGWIAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 SFLLDDIFAAMWAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 SFLLDDIFAAMWAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 HAKWCISQKGTNRTWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 HAKWCISQKGTNRTWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
Q9WZ79 PRELIMINARY; PRT; 361 AA.
AC Q9WZ79
DT 01-NAR-2002 (TrEMBLrel. 20, Created)
DT 01-NAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Endonuclease DLAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21556924; PubMed=11700027;
RA Shikawa D., Tanuma S.;
RT "Isolation and Characterization of the DLAD/DIAD Genes, Which Lie
RT Head-to-Head with the Genes for Urate Oxidase.";
RL Biochem. Biophys. Res. Commun. 288:1119-1128(2001).
DR EMBL; AF334602; AAL34449.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 361 AA; 41712 MW; C6FDD3F58F62CAC0 CRC64;

Query Match 67.1%; Score 1293.5; DB 4; Length 361;
Best Local Similarity 66.4%; Pred. No. 2.1e-109;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLISLFFALSGVLGTPETISCRNEYGEAVDWFYKLPKRTSKASBEAGLQYL 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 MWARLLRTSFALLFLGLFGLGATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 YLDSTRQTWNKSLYLINSTRSALGRITLQHLIDYTHNS--TNDTAYLIYNDGVPKPNYSRKY 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 YLDSTRSWEKSEQLMNDYKSLVGRITLQQLYEAAYASKNNNTAYLIYNDGVPKPNYSRKY 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 GHAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 GHTKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAOGLNVHFAKSS 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 QLLVLPQNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAOGLNVHFAKSS 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 FYTDDIFTGWIAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 SFLLDDIFAAMWAQKLTLLAOTWQKKOELPNCSLPYHVYNIKSIYGVTSKYFSSROD 304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 HAKWCISQKGTNRTWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 361
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
O90759 PRELIMINARY; PRT; 375 AA.
AC O90759
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II (EC 3.1.22.1).
GN CEL1/DNASEII.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP-9;
RX MEDLINE=98325194; PubMed=9658122;
RA Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;
RT "Fowlpox virus encodes non-essential homologs of cellular alpha-SNAP,
RT PC-1 and an orphan human homolog of a secreted nematode protein.";
RL J. Virol. 72:6742-6751(1998).
DR EMBL; AJ006408; CAA07012.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
KW Hydrolase.
SQ SEQUENCE 375 AA; 43064 MW; 36C641F74BAD25F6 CRC64;

Query Match 35.9%; Score 692.5; DB 12; Length 375;

```



```
RP SEQUENCE FROM N.A.
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.P., Rock D.L.;
RL Submitted (OLCT-1999) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF198100; AA64376.1; -.
DR GO: GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO: GO:0006259; F:DNA metabolism; IEA.
DR InterPro: IPR004947; DNase II.
DR Pfam: PF03265; DNase II; 1.
SQ SEQUENCE 232 AA; 26772 MW; 14C05BB3BA79C73A CRC64;

Query Match 23.6%; Score 455; DB 12; Length 232;
Best Local Similarity 40.7%; Pred. No. 2.9e-33;
Matches 96; Conservative 39; Mismatches 85; Indels 16; Gaps 7;

QY 126 LVNN--RTQGFLLHSPKPP--PVHGVEYPTSGRRYGGTGCITFCY-SQFEIDPQLL 180
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MANNSSVTCFLLHSPRPPSPVLGNTGYGVYQSMVLCINLDYKGLTALDNTLP 60

QY 181 VLPNTYSCIPSTFFHKLIMPRMCANSSSLKIPRYLAELHSAQGLNFVHFAKSSFYT 240
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 VANNENVNC---SVTNKQLNLYHLCDNKVYTLTKNVSWMESRKGKELTFAKSKYFR 117

QY 241 DDIFTGHWIAQKXTHLLAQWKKQKQLPNCNCLPYPHYVNIKSIQVTSKYFSRSQDHSK 300
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 HDLMSAWIGTLESLLSETWQRGSMITNCSSKYHVHNKISINNVGTS-FINYDHSK 176

QY 301 WCYSI---KGSANRWTCIGDLNLSHQALRGSGFICTKNHYIYQAFHKLRYGFC 353
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 WIVSLYDKKG---WVICGIDNRSPTQRHGGGACTRNGYLPKLKELTVEIYEGC 228

RESULT 10
QYVED8 PRELIMINARY; PRT; 366 AA.
AC QYVED8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2000 (TrEMBLrel. 13, Last sequence update)
DE CG7780 protein (GH10876p).
GN DNASEII OR CG7780.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
```

```
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL: AE003720; AAF55486.1; -.
RA EMBL: AY075328; AAL68195.1; -.
DR Flybase; FBgn0000477; DNaseII.
DR GO: GO:0004531; F:deoxyribonuclease II activity; NAS.
DR InterPro: IPR004947; DNase II.
DR Pfam: PF03265; DNase II; 1.
SQ SEQUENCE 366 AA; 41340 MW; 8508BC7B86276A CRC64;

Query Match 23.0%; Score 444; DB 5; Length 366;
Best Local Similarity 31.0%; Pred. No. 5.3e-32;
Matches 107; Conservative 53; Mismatches 165; Indels 20; Gaps 7;

QY 25 EISCRNEYCEAVDWFFIYKLPK---RTKASEAGLQYLIDSTR-QTNKSLYLINSTR 80
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 22 KVSCKDEAGNDVDMWHLKPLKPHYQHNDLGKDTSGTKLYLTVTSQNYDTMQMSKFIISDPL 81

QY 81 SALGRITLQHLVTHNSTNDTAVLIYNDGVPGSVVYSRQYGHAKGLLVNRTQGFLLHSV 140
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 82 SLPAQTLNPLND---DPSHTLLAAYNDQPNQTVFS-SGHAKGVVASDQETAIWIVHSV 137

QY 141 KPFPVHVHGYEYPTSGRRYGGTGCITFCYFSQFEIDFQLLVLPNTYSCFIPSTFFHKL 200
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 KPFPITPIDYSYPTSGEQYQAQSMCLCVTLKGEDLEKVGQILVYNEPHEFYQORNLATRSDEL 197

QY 201 Y-MPRMCANSSSLKIPRYLAELHSAQGLNFVHFAKSSFYTTDDIFGHWIAQKXTHLLA 258
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 FFSLEKALHGQWRTPSPFQKDLVSLDGGKFLFGKGRANVELYADVVPATLDVSLFV 257

QY 259 QTWQKKKQELPNSCLPYPHYVNIKSIQVTSKY-FSSRODHSGKCVS-----IKGS 308
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 EAWRDGAGNLPNSCKSKDKLVNVEISNPELSVDFVFTQDHSKNAVSRFTGLIYHWRVG 317

QY 309 ANRWTCIGDLNLSHQALRGSGFICTKNHYIYQAFHKLRYGFC 353
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 318 GGDWICVGINRQEQQLHGGCGTCHKSARVENLYRQLVTNYDKC 362

RESULT 11
QYBUL8 PRELIMINARY; PRT; 338 AA.
AC QYBUL8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Newborn larvae-specific protein SSI.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
ON NCBI_TaxID=6334;
```


This Page Blank (uspto)

181 AAAAGACAAAACAGGAAGTGGAGAGACTGGGTAGAGTACCTGTACCTAGACTCTACA 240
 181 AAAAGACAAAACAGGAAGTGGAGAGACTGGGTAGAGTACCTGTACCTAGACTCTACA 240
 241 ACTAGAGCTGGAGAGAGTGGAGAGCTAATGAATGACACCAAGAGTGTGTTGGGAAGG 300
 241 ACTAGAGCTGGAGAGAGTGGAGAGCTAATGAATGACACCAAGAGTGTGTTGGGAAGG 300
 301 ACATTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGCCTATCTAATA 360
 301 ACATTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGCCTATCTAATA 360
 361 TACATGATGAGTCCCTTAACCTGTGAATTAACAGTAAAGTATGACACCAAGAGT 420
 361 TACATGATGAGTCCCTTAACCTGTGAATTAACAGTAAAGTATGACACCAAGAGT 420
 421 TTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 421 TTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAAATGACAAAGTGGC 540
 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAAATGACAAAGTGGC 540
 541 ATCTGCATTAACCTTTCAAGTACCAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600
 541 ATCTGCATTAACCTTTCAAGTACCAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600
 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660
 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660
 661 CCCAGCTGTGACACAGGAGCTCATCAGAGATTCCTGGCAGGCTCTCACCACAT 720
 661 CCCAGCTGTGACACAGGAGCTCATCAGAGATTCCTGGCAGGCTCTCACCACAT 720
 721 CAGTCGGCCAGGACAAAATTCCTCATTTTGCAGAGTCCGATTCCTTTCTTGCAGAC 780
 721 CAGTCGGCCAGGACAAAATTCCTCATTTTGCAGAGTCCGATTCCTTTCTTGCAGAC 780
 781 ATCTTTGACGCTGTGATGGCTCAACGGCTGAAGACACACTTGTAAACAGAAACCTGGCAG 840
 781 ATCTTTGACGCTGTGATGGCTCAACGGCTGAAGACACACTTGTAAACAGAAACCTGGCAG 840
 841 CGAAAAACAGAGAGCTTCTTCAAACTGCTCCCTCTTACCATCTCTACAAATATAAA 900
 841 CGAAAAACAGAGAGCTTCTTCAAACTGCTCCCTCTTACCATCTCTACAAATATAAA 900
 901 GCATTAATATACGACACTCTTATTTTCAAGTCTTATCAAGATCAAGCAAGTGGTGT 960
 901 GCATTAATATACGACACTCTTATTTTCAAGTCTTATCAAGATCAAGCAAGTGGTGT 960
 961 ATTTTCCCAAGGACCAAAAATCGCTGGACATGTATTTGGAGACTTAATCGGAGTCA 1020
 961 ATTTTCCCAAGGACCAAAAATCGCTGGACATGTATTTGGAGACTTAATCGGAGTCA 1020
 1021 CACCAAGCTTCAGAGTGGAGATTCATTTGATCCAGAAATGGCAATTTACCAAGCA 1080
 1021 CACCAAGCTTCAGAGTGGAGATTCATTTGATCCAGAAATGGCAATTTACCAAGCA 1080
 1081 TTTTCAAGGATAGTATTAATATATGAAAGCTGTAAAGTAACTTGGTGAAGGACACAGT 1140
 1081 TTTTCAAGGATAGTATTAATATATGAAAGCTGTAAAGTAACTTGGTGAAGGACACAGT 1140
 1141 ACTATCAATGAAAACCTTGACAAATGGTCTTCTTCAATPACCTTTTATATTATAA 1200
 1141 ACTATCAATGAAAACCTTGACAAATGGTCTTCTTCAATPACCTTTTATATTATAA 1200
 1201 GGCCTGTGAATATATTAATCACTGATATCAAAAATTAACATATTTCTCTCATGTTT 1260
 1201 GGCCTGTGAATATATTAATCACTGATATCAAAAATTAACATATTTCTCTCATGTTT 1260
 1261 ACCATTATA 1268

Db 1261 ACCATTATA 1268
 RESULT 2
 US-10-240-709-3
 ; Sequence 3, Application US/10240709
 ; Publication No. US20030212023A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan R.
 ; APPLICANT: Trustees of Dartmouth College
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
 ; FILE REFERENCE: DC-0154
 ; CURRENT APPLICATION NUMBER: US/10/240,709
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: 09/541,840
 ; PRIOR FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: 09/574,942
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 1268
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-240-709-3
 Query Match 100.0%; Score 1268; DB 13; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGGAAAGTCTCTGCTGGCATGAAATAAATGAAACAGAAAATGATGGCAGACTG 60
 Db 1 ATGGGGAAAGTCTCTGCTGGCATGAAATAAATGAAACAGAAAATGATGGCAGACTG 60
 QY 61 CTAAAGACATCTTTGCTTCTCTTCTGGCTCTTTGGGGTGTCTGGGGGAGCAACA 120
 Db 61 CTAAAGACATCTTTGCTTCTCTTCTGGCTCTTTGGGGTGTCTGGGGGAGCAACA 120
 QY 121 ATTTTCAGCAATGAAGAGGAAAGCTCTGGACTGGTTTACTTTTATAAGTTACT 180
 Db 121 ATTTTCAGCAATGAAGAGGAAAGCTCTGGACTGGTTTACTTTTATAAGTTACT 180
 QY 181 AAAAGACAAAACAGGAAGTGGAGAGCTGGGTAGAGTACCTGTACCTAGACTCTACA 240
 Db 181 AAAAGACAAAACAGGAAGTGGAGAGCTGGGTAGAGTACCTGTACCTAGACTCTACA 240
 QY 241 ACTAGAGCTGGAGAGAGTGGAGCAACTAATGAATGACACCAAGAGTGTGTTGGGAAGG 300
 Db 241 ACTAGAGCTGGAGAGAGTGGAGCAACTAATGAATGACACCAAGAGTGTGTTGGGAAGG 300
 QY 301 ACATTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGCCTATCTAATA 360
 Db 301 ACATTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGCCTATCTAATA 360
 QY 361 TACATGATGAGTCCCTTAACCTGTGAATTAACAGTAAAGTATGACACCAAGAGT 420
 Db 361 TACATGATGAGTCCCTTAACCTGTGAATTAACAGTAAAGTATGACACCAAGAGT 420
 QY 421 TTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 Db 421 TTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 QY 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAAATGACAAAGTGGC 540
 Db 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAAATGACAAAGTGGC 540
 QY 541 ATCTGCATTAACCTTTCAAGTACCAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600
 Db 541 ATCTGCATTAACCTTTCAAGTACCAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600
 QY 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660

```

Db 601 AACCCCAACGCTAGCTGTCTCCATCCAGCCACTTCCAGGAGCTATTACATG 660
QY 661 CCCAGAGCTGTCCAGCCAGGCTCATCAGAGATTCCTGCGAGCTCTCCACACACTT 720
Db 661 CCCAGAGCTGTCCAGCCAGGCTCATCAGAGATTCCTGCGAGCTCTCCACACACTT 720
QY 721 CAGTCGGCCAGGAGCAAAATTCCTCCATTTTGCAAGTCGGATTCCTTTCTGACAC 780
Db 721 CAGTCGGCCAGGAGCAAAATTCCTCCATTTTGCAAGTCGGATTCCTTTCTGACAC 780
QY 781 ATCTTTGAGCCTGATGCTCAACGGCTGAAGCACACTTGTGTAACAGAAACCTGGCAG 840
Db 781 ATCTTTGAGCCTGATGCTCAACGGCTGAAGCACACTTGTGTAACAGAAACCTGGCAG 840
QY 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTACCAATGCTTACCAATATAAA 900
Db 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTACCAATGCTTACCAATATAAA 900
QY 901 GCAATTAATATACAGACACTTCTTATTTTCAAGTCTTATCAAGATCAGCCCAAGTGT 960
Db 901 GCAATTAATATACAGACACTTCTTATTTTCAAGTCTTATCAAGATCAGCCCAAGTGT 960
QY 961 ATTTCCAAAAGGCGACCAAAATTCGCTGAGATGTGAGACCTTAAATCGAGTCCA 1020
Db 961 ATTTCCAAAAGGCGACCAAAATTCGCTGAGATGTGAGACCTTAAATCGAGTCCA 1020
QY 1021 CACCAAGCCTTCAGAGCTGAGGATTCATTTGATCCAGATTCGCTGAGAAAGGACACAGT 1080
Db 1021 CACCAAGCCTTCAGAGCTGAGGATTCATTTGATCCAGATTCGCTGAGAAAGGACACAGT 1080
QY 1081 TTTCAAGGATTAATATATATGAAAGCTGTAAGTAAATCTTGGTGAAAGGACACAGT 1140
Db 1081 TTTCAAGGATTAATATATATGAAAGCTGTAAGTAAATCTTGGTGAAAGGACACAGT 1140
QY 1141 ACTATCAATTGAAACCTTGCAATGGTCTTCTTCCATACACCTTCTTATATTTTAA 1200
Db 1141 ACTATCAATTGAAACCTTGCAATGGTCTTCTTCCATACACCTTCTTATATTTTAA 1200
QY 1201 GGCTGTGAATATATATATGAAAGCTGTAAGTAAATCTTGGTGAAAGGACACAGT 1260
Db 1201 GGCTGTGAATATATATATGAAAGCTGTAAGTAAATCTTGGTGAAAGGACACAGT 1260
QY 1261 ACCAATTA 1268
Db 1261 ACCAATTA 1268

```

```

RESULT 3
US-10-790-589-3
; Sequence 3, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-790-589-3

```

Query Match 100.0%; Score 1268; DB 17; Length 1268;

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGAAGTGTCTCTGTGGCATGAAATAAATAAGAAACAGAAATGATGCGAAGCTG 60
Db 1 ATGGGGAAGTGTCTCTGTGGCATGAAATAAATAAGAAACAGAAATGATGCGAAGCTG 60
QY 61 CTAAGAACATCTTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120
Db 61 CTAAGAACATCTTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120
QY 121 ATTTTCATGCAAGATGAAAGAGGCAAGCTGTGCACTGCTTACTTTTATAAGTTACCT 180
Db 121 ATTTTCATGCAAGATGAAAGAGGCAAGCTGTGCACTGCTTACTTTTATAAGTTACCT 180
QY 181 AAAAGACAAAACAAAGAAAGTGGAGAGCTGGGTGTAGAGTACCTGTACCTAGACTTACA 240
Db 181 AAAAGACAAAACAAAGAAAGTGGAGAGCTGGGTGTAGAGTACCTGTACCTAGACTTACA 240
QY 241 ACTAGAGCTGGAGAAAGAGTGGAGCAACTAATGAATGACACCAAGAGTGTTTTGGGAGG 300
Db 241 ACTAGAGCTGGAGAAAGAGTGGAGCAACTAATGAATGACACCAAGAGTGTTTTGGGAGG 300
QY 301 ACATTACACAGCTATATGAAGCATATGCTCTAAGAGTAAACACACAGCCTATCTAATA 360
Db 301 ACATTACACAGCTATATGAAGCATATGCTCTAAGAGTAAACACACAGCCTATCTAATA 360
QY 361 TACAATGATGGAGTCCCTTAAACCTGTGAATTAACAGTAGAAAGTATGACACACCAAGGT 420
Db 361 TACAATGATGGAGTCCCTTAAACCTGTGAATTAACAGTAGAAAGTATGACACACCAAGGT 420
QY 421 TTTACTGCTGGAAACAGAGTTCAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 TTTACTGCTGGAAACAGAGTTCAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CCAATTCGGAAGAGCTATGATATCCACCCACAGGAGGAGCAATGACAAAGTGGC 540
Db 481 CCAATTCGGAAGAGCTATGATATCCACCCACAGGAGGAGCAATGACAAAGTGGC 540
QY 541 ATCTGATCAACTTCAAGTACCAACAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGC 600
Db 541 ATCTGATCAACTTCAAGTACCAACAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGC 600
QY 601 AACCCCAACGCTATAGCTGTCTCCATCCAGCCACTTCCAGGAGCTCAATTCACATG 660
Db 601 AACCCCAACGCTATAGCTGTCTCCATCCAGCCACTTCCAGGAGCTCAATTCACATG 660
QY 661 CCCAGCTGTGACAGGAGGCTTCCCTTCAAAAGTCCGATTCCTTTCTTGTACGAC 780
Db 661 CCCAGCTGTGACAGGAGGCTTCCCTTCAAAAGTCCGATTCCTTTCTTGTACGAC 780
QY 781 ATCTTTGAGCCTGATGCTCAACGGCTGAAGCACACTTGTGTAACAGAAACCTGGCAG 840
Db 781 ATCTTTGAGCCTGATGCTCAACGGCTGAAGCACACTTGTGTAACAGAAACCTGGCAG 840
QY 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTACCAATGCTTACCAATATAAA 900
Db 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTACCAATGCTTACCAATATAAA 900
QY 901 GCAATTAATATACAGACACTTCTTATTTTCAAGTCTTATCAAGATCAGCCCAAGTGT 960
Db 901 GCAATTAATATACAGACACTTCTTATTTTCAAGTCTTATCAAGATCAGCCCAAGTGT 960
QY 961 ATTTCCAAAAGGCGACCAAAATTCGCTGAGATGTGAGACCTTAAATCGAGTCCA 1020
Db 961 ATTTCCAAAAGGCGACCAAAATTCGCTGAGATGTGAGACCTTAAATCGAGTCCA 1020
QY 1021 CACCAAGCCTTCAGAGCTGAGGATTCATTTGATCCAGATTCGCTGAGAAAGGACACAGT 1080

```

```

Db      1021  CACCAAGCCTTCAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAAATTTACCAAGCA 1080
QY      1081  TTTCAAGGATAGTATTATATCTATGAAGCTGTAAAGCTGTAAAGCTGTGTAAGAGGACACAGGT 1140
Db      1081  TTTCAAGGATAGTATTATATCTATGAAGCTGTAAAGCTGTAAAGCTGTGTAAGAGGACACAGGT 1140
QY      1141  ACTATCATTTGAACCTTGGCAATGGCTCTCTTCCATTAACACCTCTCTTATATATTTAAA 1200
Db      1141  ACTATCATTTGAACCTTGGCAATGGCTCTCTTCCATTAACACCTCTCTTATATATTTAAA 1200
QY      1201  GGCTGTGAATATATCTTATTAACCTGCATATACAAAAATAAAACATATTTCTCTCATGTTT 1260
Db      1201  GGCTGTGAATATATCTTATTAACCTGCATATACAAAAATAAAACATATTTCTCTCATGTTT 1260
QY      1261  ACCATTTA 1268
Db      1261  ACCATTTA 1268

RESULT 4
US-10-670-863-4
; Sequence 4, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohara, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1083)
US-10-670-863-4

Query Match      85.1%; Score 1079.6; DB 17; Length 1086;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1082; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      34  ATGAACAGAAATCATGGGAGAGTCTGCTAAGAACATCTTTGGCTTTTGTCTTCTCTTGGC 93
Db      1  ATGAACAGAAATCATGGGAGAGTCTGCTAAGAACATCTTTGGCTTTTGTCTTCTCTTGGC 60

QY      94  CTCTTTGGGCTGCTGGGAGAGCAACAATTTATGACAGAAATGAAGAGGAAAGCTGTG 153
Db      61  CTCTTTGGGCTGCTGGGAGAGCAACAATTTATGACAGAAATGAAGAGGAAAGCTGTG 120

QY      154  GACTGGTTTACTTTTATTAAGTTACTTAAAGACAAAAACAAGAAAGTGGAGAGCTGGG 213
Db      121  GACTGGTTTACTTTTATTAAGTTACTTAAAGACAAAAACAAGAAAGTGGAGAGCTGGG 180

QY      214  TTAGAGTACTCTGCTGAGTACTCTACAACTAGAAAGCTGGAGAAAGTGGAGCAACTAATG 273
Db      181  TTAGAGTACTCTGCTGAGTACTCTACAACTAGAAAGCTGGAGAAAGTGGAGCAACTAATG 240

QY      274  ATGACACCAAGAGTGTGGAGGACATTAACACAGCTATATGAACATATGCTCTCT 333
Db      241  AATGACACCAAGAGTGTGGAGGACATTAACACAGCTATATGAACATATGCTCTCT 300

QY      334  AAGAGTAAACACACAGCTATCTTAATATCAATGATGGAGTCCCTAAACCTGTGAATTAC 393
Db      301  AAGAGTAAACACACAGCTATCTTAATATCAATGATGGAGTCCCTAAACCTGTGAATTAC 360

QY      394  AGTAGAAGATATGGACACACCAAGGTTTACTCTGTGGAAACAGAGTTCAAGGGTCTGTG 453

```

```

Db      361  AGCAGAAAGTATGGACACACCAAGGTTTACTCTGTGGAAACAGAGTTCAAGGTTCTGG 420
QY      454  CTGATTCATTCATCCATCCCTCAGTTTCTCCAAATTCGGGAAGAGGCTATGATTTATCCACCC 513
Db      421  CTGATTCATTCATCCATCCCTCAGTTTCTCCAAATTCGGGAAGAGGCTATGATTTATCCACCC 480
QY      514  ACAGGAGAGCAAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACCATATGAG 573
Db      481  ACAGGAGAGCAAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACCATATGAG 540
QY      574  GCAATAGATTCATGCTCTTGGTCTGCAACCCCAACCTCTATAGCTCTCATCCACGACC 633
Db      541  GCAATAGATTCATGCTCTTGGTCTGCAACCCCAACCTCTATAGCTCTCATCCACGACC 600
QY      634  ACCTTTCCACGAGAGCTCATTTACATGCCCCAGCTGTGCACCCAGGGCCAGCTCATACGAG 693
Db      601  ACCTTTCCACGAGAGCTCATTTACATGCCCCAGCTGTGCACCCAGGGCCAGCTCATACGAG 660
QY      694  ATTCCTGGCAGGCTCTCTCACACACTTCAGTGTGGCCCAAGGACAAAAATTCCTCATTTT 753
Db      661  ATTCCTGGCAGGCTCTCTCACACACTTCAGTGTGGCCCAAGGACAAAAATTCCTCATTTT 720
QY      754  GCAAGTGGATTTCTTTCTTGACAGCATCTTTGCGAGCTGTGAGGCTCAACGGCTGAAG 813
Db      721  GCAAGTGGATTTCTTTCTTGATGGCATCTTTGGAGCTGTGATGGCTCAACGGCTGAAG 780
QY      814  ACACACTTGTTTAAACAGAAACCTGCGAGCGAAAAAGACAAAGAGCTTCTTCAAACTGTCC 873
Db      781  ACACACTTGTTTAAACAGAAACCTGCGAGCGAAAAAGACAAAGAGCTTCTTCAAACTGTCC 840
QY      874  CTTCCTTACCATGTCTACAATATAAAGCAATATAATATATCAAGACACTCTTATTTCAGT 933
Db      841  CTTCCTTACCATGTCTACAATATAAAGCAATATAATATATCAAGACACTCTTATTTCAGT 900
QY      934  TCTTATCAAGATCAGCGCAAGTGTGTATTTCCCAAGAGGCGACCAAAAACTGCTGGACA 993
Db      901  TCTTATCAAGATCAGCGCAAGTGTGTATTTCCCAAGAGGCGACCAAAAACTGCTGGACA 960
QY      994  TGTATTGGAGACCTTAAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT 1053
Db      961  TGTATTGGAGACCTTAAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT 1020
QY      1054  ACCCAGAAATGGCAAAATTTACAGCAATTCAGAGCTTTCAGAAAGTGGAGGATTCATTTGT 1113
Db      1021  ACCCAGAAATGGCAAAATTTACAGCAATTTCAAGGATTTAGTATTTATCTATGAAGCTGT 1080
QY      1114  AAGTAA 1119
Db      1081  AAGTAA 1086

RESULT 5
US-10-670-863-2
; Sequence 2, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohara, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (213)..(1274)
US-10-670-863-2

Query Match
Best Local Similarity 53.1%; Score 673.8; DB 17; Length 1652;
Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3;

QY 37 AAACAGAAATGATGGCAAGACTGTCTAAGAACATCTTTGCTTGTCTCTCTCTGCGCTC 96
DB 204 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTTCTTTTGTCTCTCTCTGCGCTC 263
QY 97 TTTGGGTGCTGGGGCGAGCAACAAATTTTCATGCAAAATGAGAGGAAAGCTGTGAC 156
DB 264 TCTGGGTGCTGGGGCGAGCAACCAAGAAATCTCATGCAAAATGATGTTGAGCTGTGAC 323
QY 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAACAAAGGAAAGTGGAGACTGGGTTA 216
DB 324 TGGTTTACTTTTATAGTTTACCTTAAAGGAGTGTAGCAAGGCAAGTGAAGGCGGGCTG 383
QY 217 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGGAAAGTGTAGCAACTAATGAAT 276
DB 384 CAGTACCTGTACCTAGACTCTACAACTGGACAACTGTGACAGAGCCCTTACCTGATTAAC 443
QY 277 GACACCAAGAGTGTGGGAGGACATTAACACAGCTATATGAGCATATGCTCTTAAG 336
DB 444 AGCACCAGGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACACATAATTCACG 503
QY 337 AGTAAACACACAGCTATCTAATATACATGATGAGTCCCTTAACTGTGAATTAAGT 396
DB 504 AATGACA--CAGCTATCTAATATACAGATGCTCTGCTGCTGCTGCTGCTGCTGCTG 560
QY 397 AGAAGTATGACACACAAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGCTG 456
DB 621 ATACACTCTGTCTCCAAAGTTTCCCAAGTTC-----ATGGCTATGATGATACCACTCG 674
QY 517 GCGAGAGAAATGACAAAGTGCATCTGCAATCTTCAAGTACAAACAGTATGAGGCA 576
DB 675 GGGAGCGATATGACAAACCGGCTATGATCACTTTCCGATACAGCCAGTTTGAAGAA 734
QY 577 ATAGATTTCTCAGCTCTTGTGCTGCAACCCCAAGTCTATAGCTGCTCCATCCAGCCACC 636
DB 735 ATAGATTTCTCAGCTCTTGTGCTTACACCAACATCTAAGCTGCTTCAATCCAGGACC 794
QY 637 TTTACACAGAGCTATTCATGCCCCAGCTGTGACCCAGGCGCCAGCTCATCAGAGATT 696
DB 795 TTTCACTGGAACTTATCTACATGCCCGGATGTGTGCCAACTCCAGTTCTTAAAGATC 854
QY 697 CTGGCAGAGCTCTACACACACTTCTAGTGGCCAGGACCAAAATTTCTTCAATTTGCA 756
DB 855 CTGTCCGTTACTCTGCTGAACTGCACTCAGCCAGGCTTAAACCTTCTGCTTCAATTTGCA 914
QY 757 AAGTCGAGTTCTTTCTTGAACACATCTTTGACGCTGTGAGCTGCTCAACGGCTGAGACA 816
DB 915 AAATCAAGTTTTTATACATGATGATCTTTA CAGGATGATGATGCTCAAAAGTTGAGACA 974
QY 817 CACTTTGTTAAGAACTCTGGAGGAAAGAAAGAGCTTCTTCAAACTGCTCCCTT 876
DB 975 CATTTGTTAGCAAAACCTGGCAGAAAGAAAGAGCTTCTTCAAACTGCTCCCTG 1034
QY 877 CTTTACCATGCTCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 936
DB 1035 CTTTACCATGCTCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1094
QY 937 TATCAAGATCAGCCCAAGTGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGCAATG 996
DB 1095 CGCCAGACCATTCAAATGTTGTGTTTCCATAAAGGGCTCGCAAAATCGCTGCAATG 1154
QY 997 ATTGGAGACCTAAATGGGAGTCCACCAAGGCTTTCAGAAAGGAGGATTCATTTGTACC 1056

Db 1155 ATTGGAGACCTAAATCGAAGCCTACCAAGCCTTAAGAGGTGGAGGATTCATCTGTACA 1214
QY 1057 CAGAAATTGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTATATCTATGAAAGCTGTAA 1116
DB 1215 AAGAACTACTACATTTACCAGGCAATTTCAATAATATATATCTCCGTTATGGGTTCTGTAA 1274
QY 1117 TAAACTTGGTGAAGAGACACAGGTATCTATCATTTGAAACCTTTGACAAATGGGTCTTTCTCC 1176
DB 1275 TAAACTCGGTGAAGAGGCGACACACCTCTGTCTTGTGAAACACTGGCAGTGGAACTCTGCG 1334
QY 1177 ATTACACC--TTCTTTATATTTTAAAGGCTGTGAAT 1211
DB 1335 CTGGAATCTGTTCTCCATAATTTCAAGGCTTCTGAGT 1371

RESULT 6
US-09-949-434-1
; Sequence 1, Application US/09949434
; Patent No. US20020028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Kriesler, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-949-434-1

Query Match
Best Local Similarity 53.1%; Score 672.8; DB 9; Length 1224;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 37 AAACAGAAATGATGGCAAGACTGTCTAAGAACATCTTTGCTTGTCTCTCTCTGCGCTC 96
DB 31 AGATAGAAATGACAGCAAGCCTCTAAGAACAGTTCTTTTGTCTCTCTCTGCGCTC 90
QY 97 TTTGGGTGCTGGGGCGAGCAACAAATTTTCATGCAAAATGAGAGGAAAGCTGTGAC 156
DB 91 TCTGGGTGCTGGGGCGAGCAACCAAGATCTCATGCAAAATGATGTTGAGCTGTGAC 150
QY 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAACAAAGGAAAGTGGAGACTGGGTTA 216
DB 151 TGGTTTACTTTTATAGTTTACCTTAAAGGACTAGCAAGGCAAGTGAAGAGCGGGCTG 210
QY 217 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGAGAGTGGACCACTAATGAAT 276
DB 211 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGAGAGCTCTACCTGATTAAC 270
QY 277 GACCAACAGAGTGTGTTGGGAGGACATTAACAACTATATGAAAGCATATGCTCTTAAG 336
DB 271 AGCACCAGGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACACATTAATCCAG 330
QY 337 AGTAAACACACAGCTTATCTAATATACATGATGAGTCCCTTAAACCTGTGTAATTAAGT 396
DB 331 AATG--ACACAGCTTATCTAATATACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
QY 397 AGAAGATGAGACACACCAAGGTTTACTGCTGTGAGACAGAGTTCAGGGTTCTGGCTG 456
DB 388 AGACAGTATGAGACATGCAAGGCTGCTGCTGATGAGAAACAGAACGAGGCTTCTGGCTG 447
QY 457 ATTCAATTCCTCTCAGTTTCTTCTCAATTCGGAAGAGGCTATGATTAATCCACCA 516
DB 448 ATACACTCTGTTCCCAAGTTTCCCAAGTTC-----ATGGCTATGAGTACCAACCTCG 501
```



```

Db      1042 AAGATACCTCACTTACCGGCAATTCATATAATATATATCTCCGTATATGGTCTCTGAAG 1101
Qy      1117 TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAAAACCTTGACAAATGGTCTCTCTTCC 1176
Db      1102 TAAACTCGTGAAGGCCACACCCCTCTCTGTTGAAAAAACTGGCACCTGGAACTCTCGC 1161
Qy      1177 ATTACACCTTCTTTATATTTTA 1198
Db      1162 CTGGATCTGTCTCCATAATA 1183

```

RESULT 8

```

US-10-790-589-1
; Sequence 1, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-10-790-589-1

```

```

Query Match      53.1%; Score 672.8; DB 17; Length 1224;
Best Local Similarity 75.0%; Pred. No. 1.8e-192;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

Qy      37  AAACAGAAATGATGCGAAGCTGCTAAGAACATCCTTTGCTTTGCTTCTCTCTTGGCTC 96
Db      31  AGATAGAAATGACAGAAAGCTCTTAAGAACAGTTCTTTCTTTGCTTCTCTTTGGCTC 90
Qy      97  TTTGGGTGCTGGGGCAGCAACAATTTTCATGAGAAATGAAGAGGAAAGCTGTGGAC 156
Db      91  TCTGGGTCTGGGGACACCAGAAATCTCATGAGAAATGAATATGCTGAGCTGTGGAC 150
Qy      157  TGGTTTACTTTTATAGTTTACCTTAAAGACAAAACAAGAAAGTGGAGAGACTGGGTTA 216
Db      151  TGGTTTATCTTTTATAAGTTTACCCAAAAGGACTAGCAAGGCAAGTGAAGGGCGGGCTG 210
Qy      217  GAGTACTCTGACTGACTCTCAACTAGAGCTGGAGAGAGTGAAGTGAAGTGAAT 276
Db      211  CAGTACTCTGACTGACTCTCAAGACAACTGGAAACAGAGCTCTTACCTGATTAA 270
Qy      277  GACACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATACAGCAATATGCTCTAAG 336
Db      271  AGCACCAAGAGTGTCTGGGAGGACTTACAGCATCTGTATGACACATATTCACG 330
Qy      337  AGTAAACACACAGCTATCTATATACATATGATGGAGTCCCTTAAACCTGTGAATTACGT 396
Db      331  AATG---ACACAGCTATCTAATATACACAGTGTGCTCCCTGGATCTGTGAATTACAG 387
Qy      397  AGAAGATGAGCACACCAAGGTTTACTGCTGTGGAACAGAGTTCAGGGTCTGGCTG 456
Db      388  AGACAGTATGGAATGCCAAGGTTCTGCTGGTATGGAACAGAACCGAGGGTCTGGCTG 447
Qy      457  ATTCAATCCATCCCTCAGTTTCTCTCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516
Db      448  ATACACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGATGATCCCAACCTG 501
Qy      517  GGGAGACGAATGGACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACCAAGTATGAGGCA 576

```

```

Db      502  GGGAGCGATATGAGAAACCGGCATCTGCATCATCTTCGGATACAGCCAGTTTGAGGA 561
Qy      577  ATAGATTCAGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCTCCATCCAGCCACC 636
Db      562  ATAGATTTTCACTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCAITCCAGCACC 621
Qy      637  TTTCCACAGGAGCTCATTCACATGCCCGAGCTGTGCGACAGGGCCAGCTCATACAGATT 696
Db      622  TTTCACTGGAAACTTATCTACATGCCCGGAGTGTGCCAACTCCAGTTCCTTAAAGATC 681
Qy      697  CCTGGCAGGCTCTCTCACACACTTCAGTCGSCCCAGGAGCAAAAATTCCTCCATTTGCA 756
Db      682  CCTGTCGGTACCTTCGTGAACTGCATCAGCCAGGCTCTAAACTTCGTCCATTTGCA 741
Qy      757  AAGTCGAGTCTTTTCTTGAAGCATCTTTGACGCTGGATGGCTCAACGGCTCAAGACA 816
Db      742  AAATCAAGTCTTTTATCTGATGACATCTTTACAGATGGATAGCTCAAAGTTGAGACA 801
Qy      817  CACTTGTAAAGAACTCTGGCAGGAAAGAAAGACAGAGCTTCTTCAAACCTGCTCCCT 876
Db      802  CATTGTTAGCACAAACCTGGCAGAAAGAAACAGAGCTTCTTCAAACCTGTTCCCTG 861
Qy      877  CTTTACCATCTTACAAATATAAAGCAATTAATTTATCAAGACTCTTATTCAGTTCT 936
Db      862  CTTTACCATCTTACAAATATAAAGCAATTAATTTATCAAGACTCTTATTCAGTTCT 921
Qy      937  TATCAAGATCACGCCAAGTGTGTATTTCCTCAAAAGGCGACCAAAAATTCCTGGACATGT 996
Db      922  CGCAAGACCATTTCCAAATGTGTGTTTCCATAAAGGGCTCCGCAATTCCTGGACCTGC 981
Qy      997  ATTGGAGACCTAAATCGGAGTCCACCAAGCTTCAAGAGCTTCAAGAGTGGAGATTCAITTTGATCC 1056
Db      982  ATTGGAGACCTAAATCGAAGCTTACCAAGCTTACCAAGCTTCAAGAGTGGAGATTCAITTTGATCC 1041
Qy      1057  CAGAATTTGGCAAAATTTACCAAGCATTTCAAGGATTTAGTATTATATATGAAAGCTGTAAAG 1116
Db      1042  AAGAACTACTACATTTTACCAAGCATTTCAAGGATTTATATATGAAAGCTGTAAAG 1101
Qy      1117  TAAACTTGTGAAAGGACACAGGCTACTATCATTTGAAAAACCTTGACAAATGGTCTTCTTCC 1176
Db      1102  TAAACTCGTGAAGGCCACACCCCTCTGCTCTTGAAGAACTGGCACTGGCAATCTCGC 1161
Qy      1177  ATTACACCTTCTTTATATTTTA 1198
Db      1162  CTGGATCTGTCTCCATAATA 1183

```

RESULT 9

```

US-09-918-995-7149
; Sequence 7149, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7149
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-7149

```


Query Match 24.5%; Score 310.8; DB 10; Length 480;
Best Local Similarity 99.4%; Pred. No. 4.6e-83;
Matches 312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 335 AGAGTAAACACACAGCCTATCTATATACAAATGATGAGTCCCTAAACCTGTGAATTACA 394
Db 166 AGAGTAAACACACAGCCTATCTATATACAAATGATGAGTCCCTAAACCTGTGAATTACA 225

QY 395 GTAGAAAGTATGGACACACAAAGGTTTACTGCTGTGGAAACAGAGTTCAAGGGTTCGGC 454
Db 226 GCAGAAAGTATGGACACACAAAGGTTTACTGCTGTGGAAACAGAGTTCAAGGGTTCGGC 285

QY 455 TGATTCATTCATCCCTCAGTTTCTCCTCAATTCGGAAAGAGCTATGATATCCACCCA 514
Db 286 TGATTCATTCATCCCTCAGTTTCTCCTCAATTCGGAAAGAGCTATGATATCCACCCA 345

QY 515 CAGGAGACCAAGTGGACAAAGTGGCATCTGCATPAACTTTCAAGTACAAACAGATATGAGG 574
Db 346 CAGGAGACCAAGTGGACAAAGTGGCATCTGCATPAACTTTCAAGTACAAACAGATATGAGG 405

QY 575 CAATAGATTTCTAGCTCTTGGTCTGCAACCCCAACGCTATAGCTGCTCCATCCACCCA 634
Db 406 CAATAGATTTCTAGCTCTTGGTCTGCAACCCCAACGCTATAGCTGCTCCATCCACCCA 465

QY 635 CTTTCCACCGAG 648
Db 466 CTTTCCACCGAG 479

RESULT 10

US-10-408-167A-1
; Sequence 1, Application US/10408167A
; Publication No. US20030219428A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
Baron, Will F.
TITLE OF INVENTION: Human DNase II
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/408,167A
FILING DATE: 04-Apr-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/861034
FILING DATE: 18-May-2001
APPLICATION NUMBER: 08/639294
FILING DATE: 25-Apr-1996

ATTORNEY/AGENT INFORMATION:

NAME: Evans, David W.
REGISTRATION NUMBER: NONE
REFERENCE/DOCKET NUMBER: P1024D1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1739
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-408-167A-1

Query Match 6.5%; Score 82.4; DB 16; Length 1575;

Best Local Similarity 47.4%; Pred. No. 1.6e-13;

Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;

QY 142 GCGAAGCTGTGACTGGTTTACTTTTATAAGTT---ACCTAAAGACAAAAACAAGGAA 198
Db 155 GGGCAGCTGTAGACTTGGTTCGTGGTCTCAAGCTGCCAGCTCTTAGAGGTCGGGAG 214

QY 199 AGTGGAGAGACTGGTTAGAGTACCTGTACTTACCTAGACTCTCAACATAGAGCTGGAGGA 258
Db 215 GCGGCGCAGAGAGGGCTGCAGTATCAAGTATCTGCACGAGAGCTCCGGAGGCTCGCGGAC 274

QY 259 AGTGAGCACTAATGATGACACCAAGAGTGTTTGGGAAGGACATTAACACAGCTATAT 318
Db 275 GGCAGGGGCACTCATCAACACGCCGGGGGGCGGTGGCCGAGCTCGACGCCCTGTAC 334

QY 319 GAAGCATATGCTCTAAGAGTAAACACACAGCCCTATCTAATATACAAATGATGGAGTCCCT 378
Db 335 CGG-----AGCAACACAGCCAGCTCGCTTCTGCTCTTACAAATGACCAACCGCCT 385

QY 379 AAACC-----TGTGAATTACAGTAGAAAGTATGGACACACCAAGGTTTACTGCTGTGG 432
Db 386 CAACCCAGCAAGGCTCAGGACTCTTCCATGCGTGGGCACACGAAGGTTGCTCTGCTCTT 445

QY 433 AACAGAGTTCAAGGGTTCTGGCTGATTCATTCCTCATCTCCTCAGTTTCTTCCAAATCCGGAA 492
Db 446 GACCACGATGGGGGCTTCTGGCTGTCCACAGTGTACTTAATCTTCCCTCCACCGCCCTCC 505

QY 493 ---GAAGGCTATGATTATTCACCCACAGGGGAGACGAAATGACAAAGTGGCAATCTGCATA 549
Db 506 TCTGCTGCATACAGCTGGGCTCATAGGCTGTACTACGGGCGAGACCTGCTGTGTGTG 565

QY 550 ACTTTCAGATACACACAGTATGAGCAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAC 609
Db 566 TCTTTTCCCTTCGCTCAGTTCTCGAAGATGGGCAAGCAGCTGACCTACCTACCCCTGG 625

QY 610 GTCTATAGCTGCTCCATCCCGACCACTTTTCAACAGGAGCTCANTCATATGCCAGCTG 669
Db 626 GTCTATACCTACAGCTGGAAGGGATCTTTGCCACGAGATTTCCCGACTTGGAGATGTG 685

QY 670 TGCACAGGGGCGAGCTCATCAGAGATTCCTGGCAGGCTCCTCAACACTTCAGTGGGCC 729
Db 686 GTCAGGGCCACACAGCTTAGCCAAAGACCTGGAACAGCAGCATCACACTCACATCCAG 745

QY 730 CAGGACAAAAATTCCTCCATTTTGCAGAGTGGATTCTTTTCTTGACGACATCTTTGCA 789
Db 746 GCGGGGCTGTTTTCAGAGCTTTTCCAAAGTTTCCAAATTTGGAGATGACCTGTACTCC 805

QY 790 GCCTGGATGGCTCAACGGCTGAAGACACACTTGTTAACAGAAACCTGGCAGCGAAAAAGA 849
Db 806 GGTGTTGGCAGCAGCGCTTGGTACCAACCTGAGGTCAGTTCTGGCACAACACTGTA 865

QY 850 CAAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCAATGTCTACAAATATAAAGCAATAAA 909
Db 866 GGCATCTGCTCTTAACCTGCTCGATATCTGGCAGGTCTCTGAATGTGAACCCAGATAGCT 925

QY 910 TTATCAACGACA-----CTCTTATTTTCACTTATCAGATCAGCAGCAAGTGGTGTATT 963
Db 926 TTCCCTGGACACAGCGCGCCCAAGCTTCAACAGCAGAGGACCATCTCAAAATTTGGTGG 985

QY 964 TCCCAAAAGGGGCAACAAAAATCGCTGGACATGATTTGGAGACCTAAATCGGAGTCCACAC 1023
Db 986 TCCCAAAAGGGGCC-----CTGGACCTCGCTGGGTGACATGAATCGGAACCAAGGA 1036

QY 1024 CAAGCCTTTCAGAAAGTGGAGATTCTAATTTGTACCCAGAAATGGCAAAATTTACAGCATTT 1083
Db 1037 GAGGAGCAACGGGTTGGGGGCACACTGTGTGCCAGCTCCAGCCCTCTTGGAAAGCCTTC 1096

QY 1084 CAAGGATTTAGTATTATTACTATGAAAGCTGTAA 1115

Db 1097 CAGCGCTGTGAAGAACTACCAGCCCTGTAA 1128

RESULT 11

US-09-908-975-13222

; Sequence 13222, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi

; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli

; APPLICANT: MINTZ, Liat

; APPLICANT: FAIGLER, Simchon

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE

; FILE REFERENCE: 36688-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13222

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-908-975-13222

Query Match 4.7%; Score 60; DB 10; Length 60;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTGGACATGTAATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGAT 1045

Db 1 GCTGGACATGTAATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGAT 60

RESULT 12

US-09-908-975-2418

; Sequence 2418, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi

; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli

; APPLICANT: MINTZ, Liat

; APPLICANT: FAIGLER, Simchon

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE

; FILE REFERENCE: 36688-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2418

; LENGTH: 65

; TYPE: DNA

; ORGANISM: Rattus norvegicus

US-09-908-975-2418

Query Match 4.5%; Score 57; DB 10; Length 65;

Best Local Similarity 92.3%; Pred. No. 9.3e-07;

Matches 60; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 984 TCGCTGGACATGTAATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGG 1043

Db 1 TCGCTGGACCTGTAATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGG 60

QY 1044 ATTCA 1048

Db 61 ATTTA 65

RESULT 13

US-09-918-995-1186

; Sequence 1186, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1186

; LENGTH: 478

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(478)

; OTHER INFORMATION: n = A, T, C or G

US-09-918-995-1186

Query Match 4.0%; Score 50.6; DB 10; Length 478;

Best Local Similarity 45.5%; Pred. No. 0.00031;

Matches 179; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 527 ATGCGACAAAGTGGCATCTGCATTAACCTTTCAAGTACACCAAGTATGAGGCAATGATTC 586

Db 35 ACGGGCAGACCTGCTCTGTGTCTCTTTCCCTTGGCTCAGTCTCTGAAGATGGGCAAGC 94

QY 587 AGCTCTGGTCTGCAACCCCAACGCTCTATAGCTGTCTCCATCCAGCCAGCCCTTTCCAGG 646

Db 95 AGCTGACCTACACTACCTACCTCCCTGGGTCTATACTACCACTGGAGGGATCTTTGCCAGG 154

QY 647 AGCTCAATTCATGCCCCAGCTGTGCACAGGCCAGCTCATCAGAGTTCTCTGGGAGGC 706

Db 155 AATTCCTCCGACTTGGAGAATGTGTCAAGGGCCACCACTGTTAGCCCAAGAACCTTGAACA 214

QY 707 TCCTCACCACATCTCAGTGGCCCGCCAGGGACAAATAATTCCTCCATTTTGCAGGCTGATT 766

Db 215 GCAGCATCAGCTCAGATCCAGCCGGGCTGTTTCAGAGCTTTGCCAGTTCCAGCA 274

QY 767 CTCTTCTTGCAGCATCTTTTGGAGCTTGGATGGCTCAACGGCTGGAAGACACACTTGTAA 826

Db 275 AATTTGGAGATGACCTGTACTCCGGCTGTGTGGCAGCAGCCCTTGGTACCACTTCAGG 334

QY 827 CAGAACCTTGGAGCGGAGGAGAGAGCTTCCTTCAAACTGCTCCCTTCCTTACCATTG 886

Db 335 TCCAGTTCTGGCAGCAAACTGTAGGCATCTGCGCTCTAACTGCTCGGATATCTGGCAGG 394

QY 887 TCTACAATATAAAGCAATTAATAATATACAGAC 919

Db 395 TTCTGATGTGAACAGATAGCTTTCCCTGGAC 427

RESULT 14

US-09-983-965-4982

; Sequence 4982, Application US/09983965

; Patent No. US20020137160A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

us-10-790-589-3.rnpb

FILE REFERENCE: 37-21(10297)C

! : CURRENT AFFILIATION NUMBER: 95/3
: CURRENT FILING DATE: 2001-10-26

;; PRIOR APPLICATION NUMBER: US 03/463,231
: PRIOR FILING DATE: 1999-12-15

;; PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17

; NUMBER OF SEQ ID NOS: 5912

```

; LENGTH: 222
;

```

ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 27-LIB34-

1. *Chlorophyll a* (Chl *a*)
 2. *Chlorophyll b* (Chl *b*)
 3. *Chlorophyll c* (Chl *c*)
 4. *Chlorophyll d* (Chl *d*)
 5. *Chlorophyll e* (Chl *e*)
 6. *Chlorophyll f* (Chl *f*)
 7. *Chlorophyll g* (Chl *g*)
 8. *Chlorophyll h* (Chl *h*)
 9. *Chlorophyll i* (Chl *i*)
 10. *Chlorophyll j* (Chl *j*)
 11. *Chlorophyll k* (Chl *k*)
 12. *Chlorophyll l* (Chl *l*)
 13. *Chlorophyll m* (Chl *m*)
 14. *Chlorophyll n* (Chl *n*)
 15. *Chlorophyll o* (Chl *o*)
 16. *Chlorophyll p* (Chl *p*)
 17. *Chlorophyll q* (Chl *q*)
 18. *Chlorophyll r* (Chl *r*)
 19. *Chlorophyll s* (Chl *s*)
 20. *Chlorophyll t* (Chl *t*)
 21. *Chlorophyll u* (Chl *u*)
 22. *Chlorophyll v* (Chl *v*)
 23. *Chlorophyll w* (Chl *w*)
 24. *Chlorophyll x* (Chl *x*)
 25. *Chlorophyll y* (Chl *y*)
 26. *Chlorophyll z* (Chl *z*)
 27. *Chlorophyll aa* (Chl *aa*)
 28. *Chlorophyll ab* (Chl *ab*)
 29. *Chlorophyll ac* (Chl *ac*)
 30. *Chlorophyll ad* (Chl *ad*)
 31. *Chlorophyll ae* (Chl *ae*)
 32. *Chlorophyll af* (Chl *af*)
 33. *Chlorophyll ag* (Chl *ag*)
 34. *Chlorophyll ah* (Chl *ah*)
 35. *Chlorophyll ai* (Chl *ai*)
 36. *Chlorophyll aj* (Chl *aj*)
 37. *Chlorophyll ak* (Chl *ak*)
 38. *Chlorophyll al* (Chl *al*)
 39. *Chlorophyll am* (Chl *am*)
 40. *Chlorophyll an* (Chl *an*)
 41. *Chlorophyll ao* (Chl *ao*)
 42. *Chlorophyll ap* (Chl *ap*)
 43. *Chlorophyll aq* (Chl *aq*)
 44. *Chlorophyll ar* (Chl *ar*)
 45. *Chlorophyll as* (Chl *as*)
 46. *Chlorophyll at* (Chl *at*)
 47. *Chlorophyll au* (Chl *au*)
 48. *Chlorophyll av* (Chl *av*)
 49. *Chlorophyll aw* (Chl *aw*)
 50. *Chlorophyll ax* (Chl *ax*)
 51. *Chlorophyll ay* (Chl *ay*)
 52. *Chlorophyll az* (Chl *az*)
 53. *Chlorophyll aza* (Chl *aza*)
 54. *Chlorophyll abz* (Chl *abz*)
 55. *Chlorophyll acz* (Chl *acz*)
 56. *Chlorophyll adz* (Chl *adz*)
 57. *Chlorophyll aez* (Chl *aez*)
 58. *Chlorophyll afz* (Chl *afz*)
 59. *Chlorophyll agz* (Chl *agz*)
 60. *Chlorophyll ahz* (Chl *ahz*)
 61. *Chlorophyll aiz* (Chl *aiz*)
 62. *Chlorophyll ajz* (Chl *ajz*)
 63. *Chlorophyll akz* (Chl *akz*)
 64. *Chlorophyll alz* (Chl *alz*)
 65. *Chlorophyll amz* (Chl *amz*)
 66. *Chlorophyll anz* (Chl *anz*)
 67. *Chlorophyll aoz* (Chl *aoz*)
 68. *Chlorophyll apz* (Chl *apz*)
 69. *Chlorophyll aqz* (Chl *aqz*)
 70. *Chlorophyll arz* (Chl *arz*)
 71. *Chlorophyll asz* (Chl *asz*)
 72. *Chlorophyll atz* (Chl *atz*)
 73. *Chlorophyll auz* (Chl *auz*)
 74. *Chlorophyll avz* (Chl *avz*)
 75. *Chlorophyll awz* (Chl *awz*)
 76. *Chlorophyll axz* (Chl *axz*)
 77. *Chlorophyll ayz* (Chl *ayz*)
 78. *Chlorophyll azz* (Chl *azz*)
 79. *Chlorophyll azaa* (Chl *aza*
 80. *Chlorophyll abz* (Chl *abz*)
 81. *Chlorophyll acz* (Chl *acz*)
 82. *Chlorophyll adz* (Chl *adz*)
 83. *Chlorophyll aez* (Chl *aez*)
 84. *Chlorophyll afz* (Chl *afz*)
 85. *Chlorophyll agz* (Chl *agz*)
 86. *Chlorophyll ahz* (Chl *ahz*)
 87. *Chlorophyll aiz* (Chl *aiz*)
 88. *Chlorophyll ajz* (Chl *ajz*)
 89. *Chlorophyll akz* (Chl *akz*)
 90. *Chlorophyll alz* (Chl *alz*)
 91. *Chlorophyll amz* (Chl *amz*)
 92. *Chlorophyll anz* (Chl *anz*)
 93. *Chlorophyll aoz* (Chl *aoz*)
 94. *Chlorophyll apz* (Chl *apz*)
 95. *Chlorophyll aqz* (Chl *aqz*)
 96. *Chlorophyll arz* (Chl *arz*)
 97. *Chlorophyll asz* (Chl *asz*)
 98. *Chlorophyll atz* (Chl *atz*)
 99. *Chlorophyll auz* (Chl *auz*)
 100. *Chlorophyll avz* (Chl *avz*)
 101. *Chlorophyll awz* (Chl *awz*)
 102. *Chlorophyll axz* (Chl *axz*)
 103. *Chlorophyll ayz* (Chl *ayz*)
 104. *Chlorophyll azz* (Chl *azz*)
 105. *Chlorophyll azaa* (Chl *aza*
 106. *Chlorophyll abz* (Chl *abz*)
 107. *Chlorophyll acz* (Chl *acz*)
 108. *Chlorophyll adz* (Chl *adz*)
 109. *Chlorophyll aez* (Chl *aez*)
 110. *Chlorophyll afz* (Chl *afz*)
 111. *Chlorophyll agz* (Chl *agz*)
 112. *Chlorophyll ahz* (Chl *ahz*)
 113. *Chlorophyll aiz* (Chl *aiz*)
 114. *Chlorophyll ajz* (Chl *ajz*)
 115. *Chlorophyll akz* (Chl *akz*)
 116. *Chlorophyll alz* (Chl *alz*)
 117. *Chlorophyll amz* (Chl *amz*)
 118. *Chlorophyll anz* (Chl *anz*)
 119. *Chlorophyll aoz* (Chl *aoz*)
 120. *Chlorophyll apz* (Chl *apz*)
 121. *Chlorophyll aqz* (Chl *aqz*)
 122. *Chlorophyll arz* (Chl *arz*)
 123. *Chlorophyll asz* (Chl *asz*)
 124. *Chlorophyll atz* (Chl *atz*)
 125. *Chlorophyll auz* (Chl *auz*)
 126. *Chlorophyll avz* (Chl *avz*)
 127. *Chlorophyll awz* (Chl *awz*)
 128. *Chlorophyll axz* (Chl *axz*)
 129. *Chlorophyll ayz* (Chl *ayz*)
 130. *Chlorophyll azz* (Chl *azz*)
 131. *Chlorophyll azaa* (Chl *aza*
 132. *Chlorophyll abz* (Chl *abz*)
 133. *Chlor*

Best Local Similarity 79.7%; Pred. No

[illegible]

0 1 2 3 4 5 6 7 8 9

[illegible]

US-10-663-561-352/c

Publication No. US20040166509A1

APPLICANT: DENSILOW, NANCY D.

APPLICANT: LARRIN, FAIRICK M.
APPLICANT: FOI-MAR LEROY C.

; APPLICANT: HEMMER, MICHAEL D.
: TITLE OF INVENTION: DETECTING; FILE REFERENCE: 5853-238
; CURRENT APPLICATION NUMBER: IIS

; CURRENT FILING DATE: 2003-09-
NUMBER OF CERTS IN NOC: ECC

```

; SOFTWARE: PatentIn version 3.2
;

```

LENGTH: 646

; ORGANISM: *Cyprinodon variegatus*

Best Local Similarity 57.5%;

100

.....

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

```
Job time : 624.173 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:41:52 ; Search time 3092.15 Seconds
(without alignments)
12245.607 Million cell updates/sec

Title: US-10-790-589-3

Perfect score: 1268

Sequence: 1 atggggaaagtgcctgctg.....tctctcatgtttaccattta 1268

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	870.4	68.6	893	13	BQ681950
C 2	723.6	57.1	758	14	CD365871
C 3	707.2	55.8	729	14	CD365107
C 4	690.6	54.5	727	14	CD366106

5	681.2	53.7	923	12	BG761484
6	672.2	53.0	2689	11	AK079224
C 7	621.8	49.0	676	9	AI802154
C 8	580.4	45.8	609	12	BG149286
C 9	553	43.6	554	9	AI420898
C 10	550.2	43.4	571	10	AW295184
C 11	528	41.6	536	9	AI694455
C 12	511.4	40.3	521	9	AI820599
C 13	482.8	38.1	487	9	AI791998
C 14	467.8	36.9	482	13	BX280577
C 15	441.8	34.8	446	9	AA988125
C 16	415	32.7	448	10	BF108594
C 17	395.6	31.2	714	10	BE966684
C 18	349.2	27.5	397	13	BX482994
C 19	315	24.8	695	14	CF725003
C 20	303.4	23.9	521	13	BQ373395
C 21	282.4	22.3	887	13	BU241430
C 22	280.2	22.1	510	9	AI526873
C 23	272.4	21.5	710	13	BU240964
C 24	256.4	20.2	546	13	BX526247
C 25	227.6	17.9	432	9	AI875954
C 26	225.2	17.8	792	13	BU240892
C 27	220.8	17.4	669	12	BI392355
C 28	216	17.0	765	13	BU397346
C 29	210	16.6	258	10	BE937933
C 30	209	16.5	800	13	BU457064
C 31	207.6	16.4	543	9	AI048641
C 32	207	16.3	999	13	BU235150
C 33	203.6	16.1	643	13	BU215085
C 34	202.2	15.9	406	14	C8080400
C 35	198.2	15.6	522	13	BU299065
C 36	194	15.3	424	14	CB796507
C 37	191	15.1	473	14	CB729923
C 38	188	14.8	381	9	AI876076
C 39	187.8	14.8	223	9	AI658810
C 40	186.4	14.7	652	10	B8627834
C 41	186.2	14.7	643	13	BU280296
C 42	182	14.4	948	14	CF590064
C 43	160.4	12.6	565	12	BM426634
C 44	150.6	11.9	476	9	AI549457
C 45	146	11.5	454	12	BI281397

ALIGNMENTS

RESULT 1
BQ681950
LOCUS
DEFINITION
AGENCY: 818335 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262400
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BQ681950 893 bp mRNA linear EST 15-JUL-2002
AGENCY: 818335 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262400
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BQ681950 893 bp mRNA linear EST 15-JUL-2002
AGENCY: 818335 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262400
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES	source	Location/Qualifiers
1.	.893	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/IMAGE="IMAGE:6262400"
		/tissue_type="melanotic melanoma, cell line"
		/lab_host="DH10B (phage-resistant)"
		/clone_lib="NIH_MGC_112"
		/notes="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by King Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life technologies). Note: this is a NIH_MGC Library."
ORIGIN		
Query Match	68.6%;	Score 870.4; DB 13; Length 893;
Best Local Similarity	98.5%;	Pred. No. 4.3e-233;
Matches 877;	Conservative 0;	Mismatches 13; Indels 0; Gaps 0;
QY	365	ATGATGGAGTCCCTAAACCTGTGAATTACAGTAGAAGATATGGACACACAAAGGTTTAC 424
DB	1	ATGATGGAGTCCCTAAACCTGTGAATTACAGTAGAAGATATGGACACACAAAGGTTTAC 60
QY	425	TGCTGTGGAAACAGAGTTCAAGGGTTCTGGGTGATTCATTCCATCCCTCAGTTTCCTCCAA 484
DB	61	TGCTGTGGAAACAGAGTTCAAGGGTTCTGGGTGATTCATTCCATCCCTCAGTTTCCTCCAA 120
QY	485	TTCCGGAAGAGGCTATGATTATCCACCCACGGGAGACGAATGTACAAAGTGGCATCT 544
DB	121	TTCCGGAAGAGGCTATGATTATCCACCCACGGGAGACGAATGTACAAAGTGGCATCT 180
QY	545	GCATAACTTTCAAGTACAAACAGATATAGGCAATAGATTCTCAGCTCTTGGTCTGCAACC 604
DB	181	GCATAACTTTCAAGTACAAACAGATATAGGCAATAGATTCTCAGCTCTTGGTCTGCAACC 240
QY	605	CCAAAGTCTATAGTGTCTCATCCAGCCACTTTTACAGGAGCTCATTCATGACGCCC 664
DB	241	CCAAAGTCTATAGTGTCTCATCCAGCCACTTTTACAGGAGCTCATTCATGACGCCC 300
QY	665	AGCTGTGCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCTCCACACACTTCAGT 724
DB	301	AGCTGTGCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCTCCACACACTTCAGT 360
QY	725	CGGCCACAGGACAAAATTCCTCCATTTTGCAAAGTCGGATTCCTTTTTCAGCATCT 784
DB	361	CGGCCACAGGACAAAATTCCTCCATTTTGCAAAGTCGGATTCCTTTTTCAGCATCT 420
QY	785	TTGCAGCCTGGATGGCTCAACGGCTGAAGACACACTTGTAAACAGAACTGGCAGGAA 844
DB	421	TTGCAGCCTGGATGGCTCAACGGCTGAAGACACACTTGTAAACAGAACTGGCAGGAA 480
QY	845	AAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCTTACCATGTCTCAATATATAAAGCAA 904
DB	481	AAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCTTACCATGTCTCAATATATAAAGCAA 540
QY	905	TTAAATATACAGCACTCTTATTTTCAGTCTTATTCAGATCATCGCAAGTGGTATTT 964
DB	541	TTAAATATACAGCACTCTTATTTTCAGTCTTATTCAGATCATCGCAAGTGGTATTT 600
QY	965	CCCAAAAGGGACCAAAATTCGCTGGACATGTATGGAGACCTTAATTCGGAGTCCACAC 1024
DB	601	CCCAAAAGGGACCAAAATTCGCTGGACATGTATGGAGACCTTAATTCGGAGTCCACAC 660
QY	1025	AAGCCCTCAGAGTCGAGGATTCATTCTGTACCCAGAAATGGCAATTTTACCAAGCATTT 1084
DB	661	AAGCCCTCAGAGTCGAGGATTCATTCTGTACCCAGAAATGGCAATTTTACCAAGCATTT 720
QY	1085	AAGGATTAGTATTATATCTATGAAGCTCTAAGTAAACTTGTGTGAAGGACACAGGTACTA 1144


```

RESULT 4
CD366106/c
LOCUS
DEFINITION
  UI-H-FT1-bjs-1-13-0-UI.s1 NCI CGAP_FTI Homo sapiens cDNA clone
  UI-H-FT1-bjs-1-13-0-UI 3', mRNA sequence.
ACCESSION
  CD366106
VERSION
  CD366106.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 727)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. Gary W. Humminghake, U of I
  CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/cgap.html
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source
  1..727
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="UI-H-FT1-bjs-1-13-0-UI"
  /tissue_type="Aveolar Macrophage"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="NCI-CGAP_FTI"
  /note="Organ: Lung; Vector: pVT73-Pac (Pharmacia) with a
  modified polylinker; Site 1: EcoR I; Site 2: Not I;
  NCI CGAP FTI is a normalized cDNA library constructed from
  a pool of 81 RNA samples from Alveolar Macrophages
  challenged with different treatments. The library was
  normalized according to Bonaldo, Lennon and Soares, Genome
  Research, 6:791-806, 1996. First strand cDNA synthesis was
  primed with an oligo-dT primer containing a Not I site.
  Double stranded cDNA was ligated to an EcoR I adaptor,
  digested with Not I, and cloned directionally into
  pVT73-Pac vector. The oligonucleotide used to prime the
  synthesis of first-strand cDNA contains a library tag
  sequence that is located between the Not I site and the
  (dT)18 tail. The sequence tag for this library is
  GGCCATGCCG. The tissue was provided by Dr. Gary W.
  Humminghake of the University of Iowa.
  TAG TISSUE=Human Lung Aveolar Macrophage
  TAG LIB=UI-H-FT1
  TAG_SEQ=GGCCATGCCG"
ORIGIN
  Query Match 54.5%; Score 690.6; DB 14; Length 727;
  Best Local Similarity 99.0%; Pred. No. 1.4e-182;
  Matches 704; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
  QY 558 GTACAAACAGTAGGCAATGAGTTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAG 617
  DB 727 GTACAAACAGTAGA-GCAATGAGTTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAG 669
  QY 618 CTGCTCCATCCAGCCACTTTCACAGGAGTCTATTACATGCCGCCAGCTGGACCCAG 677
  DB 668 CTGCTCCATCCAGCCACTTTCACAGGAGTCTATTACATGCCGCCAGCTGGACCCAG 609
  QY 678 GGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACCACATTCAGTCGCCCCAGGACA 737
  DB 608 GCCAGNCTCATCAGAGATTCCTGGCAGGCTCCTCACCACATTCAGTCGCCCCAGGACA 549
  QY 738 AAAATTCCTCCATTTTGCAAGTCGGATTCCTTTCTTGACGACATCTTTTCAGCTCGAT 797
  DB 548 AAAATTCCTCCATTTTGCAAGTCGGATTCCTTTCTTGACGACATCTTTTCAGCTCGAT 489
  QY 798 GGCTCAACGGCTGAAGACACACATTTGTTAAACAGAAACCTGGCAGCGAAAGAGAGCT 857
  DB 488 GGCTCAACGGCTGAAGACACACATTTGTTAAACAGAAACCTGGCAGCGAAAGAGAGCT 429
  QY 858 TCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAAAGCAATTAATTCACG 917
  DB 428 TCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAAAGCAATTAATTCACG 369
  QY 918 ACATCTTATTTAGTTTCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAGAGGCAC 977
  DB 368 ACATCTTATTTAGTTTCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAGAGGCAC 309
  QY 978 CAAAAATCGCTGGACATGTTATTTGAGAGACTAAATCGGAGTCCACACCAAGCCTTCAGAAG 1037
  DB 308 CAAAAATCGCTGGACATGTTATTTGAGAGACTAAATCGGAGTCCACACCAAGCCTTCAGAAG 249
  QY 1038 TGGAGGATTCATTTGTATCCCGAATTTGGCAAAATTTACAGCAATTCAGAGTATTAGTATT 1097
  DB 248 TGGAGGATTCATTTGTATCCCGAATTTGGCAAAATTTACAGCAATTCAGAGTATTAGTATT 189
  QY 1098 ATACTATGAAGCTGTAAAGTAAACTTGTGTAAAGGACACACAGGTACTATCATTTGAAAACCT 1157
  DB 188 ATACTATGAAGCTGTAAAGTAAACTTGTGTAAAGGACACACAGGTACTATCATTTGAAAACCT 129
  QY 1158 TGACAAATGGGTCTTCTTCCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATCTT 1217
  DB 128 TGACAAATGGGTCTTCTTCCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATCTT 69
  QY 1218 ATAACTGTCATATCAGAAAATAAACAATATTTCTCTCATGTTTACCATTTA 1268
  DB 68 ATAACTGTCATATCAGAAAATAAACAATATTTCTCTCATGTTTACCATTTA 18
  RESULT 5
  BG761484 923 bp mRNA linear EST 15-MAY-2001
  602718694F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858894 5',
  mRNA sequence.
  BG761484
  ACCESSION
  BG761484.1
  VERSION
  BG761484.1
  KEYWORDS
  EST.
  SOURCE
  Homo sapiens (human)
  ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  1 (bases 1 to 923)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: ATCC/DCTD/DTIP
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLCM1713 row: k column: 23
  High quality sequence stop: 845.
  Location/Qualifiers
  1..923
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:4858894"
FEATURES
  source
  1..923
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:4858894"

```

```

/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 49"
/notes="Organ; skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using Zap-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN

Query Match	53.7%;	Score 681.2;	DB 12;	Length 923;
Best Local Similarity	95.9%;	Pred. 6.3e-180;		
Matches 821;	Conservative 0;	Mismatches 13;	Indels 22;	Gaps 11;
QY	367	GATGGAGTCCCTAAACCTGTGCAATTACAGTGAAGAAGTATGACACACACCAAGGTTTACTG	426	
Db	1	GATGGAGTCCCTAAACCTGTGCAATTACAGCAGAAGTATGACACACCAAGGTTTACTG	60	
QY	427	CTGTGGAACAGAGTTCGAAGGGTCTTGCTGATTCATTCCATCCCTCAGTTTCCTCCCAATT	486	
Db	61	CTGTGGAACAGAGTTCGAAGGGTCTTGCTGATTCATTCCATCCCTCAGTTTCCTCCCAATT	120	
QY	487	CCGGAAGAAGGCTATGATTTATCCACCAACAGGGAGAGAAATGACAAAAGTGGCATCTGC	546	
Db	121	CCGGAAGAAGGCTATGATTTATCCACCAACAGGGAGAGAAATGACAAAAGTGGCATCTGC	180	
QY	547	ATTAACCTTTCAAGTACACCAAGTATGAGGCAATAG-----ATTCTCAGCTCTTGG	595	
Db	181	ATTAACCTTTCAAGTACACCAAGTATGAGGCAATAGTAAACCTAAATTTCTCAGCTCTTGG	240	
QY	596	TCTGCAACCCCAACCGTCTATAGCTGCTCCATCC-AGCCACCTTTTCCACAGAGGCTCAT	653	
Db	241	TCTGCAACCCCAACCGTCTATAGCTGCTCCATCCCTAGCCACCTTTTCCACAGAGGCTCAT	300	
QY	654	TCACATGCCCCCA-GCTGTGCA-CCAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTC	711	
Db	301	TCACATGCCCCCAAGCTGTGACGCGAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTC	360	
QY	712	ACCACACTTCAGTCGGCCGAGGACAAAAATTCCTCCATTTTGCAAGTCGGAATCTTTTT	771	
Db	361	ACCACACTTCAGTCGGCCGAGGACAAAAATTCCTCCATTTTGCAAGTCGGAATCTTTTT	420	
QY	772	CTTGAAGCACATCTTTGCAAGCTGGATGGCTCAACGGCTGAAGACACACTTGTTTAAACAGAA	831	
Db	421	CTTGAAGCACATCTTTGCAAGCTGGATGGCTCAACGGCTGAAGACACACTTGTTTAAACAGAA	480	
QY	832	ACCTGGC--AGCGAAAAAGACAAGAGCTTCCTTTCAAACTGCTCCCTTCCTTACCATGTCT	889	
Db	481	ACCTGGCTAGCGAATAAAGACAAGAGCTTCCTTTCAAACTGCTCCCTTCCTTACCATGTCT	540	
QY	890	ACAATATAAAGCAATTAAATTATCAGGACACTCTTATTTTCAGTCTCTTATCAAGATCAAG	949	
Db	541	ACAATATAAAGCAATTAAATTATCAGGACACTCTTATTTTCAGTCTCTTATCAAGATCAAG	600	
QY	950	CCAAGTGGTGTATTTCCCAAAAGGGCACCAAA-AAAATCGCTGGAACATGTATTGGAGACCTA	1008	
Db	601	CCAAGTGGTGTATTTCCCAAAAGGGCACCAAGAAATCGCTGGACATGTATTGGAGACCTA	660	
QY	1009	AATCGGAGTCCACACCAAGCCTTTCAGAAGTGG-AGGATTCAATTTGTACCC-AGCAATTGGC	1066	
Db	661	AATCGGAGTCCACACCAAGCCTTTCAGAAGTGGAGGATTCATTTTGTACCAAGATTTGGC	720	
QY	1067	AAATTTACCAAGCATTTCAAGGATAGTATTTATCTATGAAAGCTGTAAAGTAACTTGGT	1126	
Db	721	AAATTTACCAAGCATTTCAAGGATAGTATTTATCTATGAAAGCTGTAAAGTAACTTGGT	780	
QY	1127	GAAAGGACACAGTACTAT-CATTGAAAACCTTGAACAATGGGCTTCTTCCATTAACCT	1185	
Db	781	GAAAGGACCAAGTACTATCCATTGAAAACCTTGAAGATGGGCTTCTTCCATTA-ACCT	839	

QY	1186	TC	T	T	A	T	A	T	T	T	A	A	G	1201
D _b	840	TC	T	T	A	T	A	T	T	G	A	A	G	855

RESULT 6	DEFINITION
AK079224	
LOCUS	

LOCUS	AK079224	2689 bp	mrna	linear	HTC 19-SEP-2003
DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530041E10 product:deoxyribonuclease II beta, full insert sequence.				

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

ORGANISM	REFERENCE
<i>Mus musculus</i>	1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
	Carninci, P. and Hayashizaki, Y.
	High-efficiency full-length cDNA cloning
	Meth. Enzymol. 303, 19-44 (1999)
	99279253
	PUBMED 10349636

REFERENCE	2
AUTHORS	Carrinici,P., Shibata,Y., Hayata,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PubMed
3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	20530913	11076861

REFERENCE
 4
AUTHORS
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE
 Functional annotation of a full-length mouse cDNA collection
JOURNAL
 Nature 409, 685-690 (2001)

REFERENCE	Nature 405, 685-690 (2001)
FORWARD	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)

6 (bases 1 to 269)
120027

REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arai, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashi, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE
Miyamoto, M. and Hayashizaki, I.
JOURNAL
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp).
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

Location/Qualifiers
1. .2689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:9530041E10"
/db_xref="MGI:2399700"
/db_xref="taxon:10090"
/clone="9530041E10"
/sex="male"
/tissue type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
335. .1399
/note="unnamed protein product; deoxyribonuclease II beta
(MGI:1913283, GB|NW_019957, evidence: BLASTN, 99%,
match=1529)
putative"
/codon_start=1
/protein_id="BAC37579.1"
/db_xref="GI:26347861"
/translation="MTAKPLRTVLISLFFALSGVLGTPEISCRNEYGEAVDFIFYKL
PKTSKSEERAGLYLDSPTWNSLYLINSALGKTLQHLVDHNSINDTAY
LIYNDGPGSVNRYOYGHAKGLVWNETQGFVLIHSPKPPVHGYEPTSGRRYGO
TGICITGYSEIDFOLLVQPNVYSCFIPSPFHKLIIYMPKNCANSSSLKIPVY
LAELHSAQGLNHFVFAKSFYDIDFTGWIAQKTHLLAFTWKKKQBELPNSCLPY
HVYNIKGISYTSKYSFSDRSHKWCVSIGKSANRWTCIGDLNLSLHQAURGGGFICT
KNHYIQAFHKLVLRYLPGCK"

ORIGIN

Query Match 53.0%; Score 672.2; DB 11; Length 2689;
Best Local Similarity 75.0%; Pred. No. 2.9e-177;
Matches 883; Conservative 0; Mismatches 283; Indels 11; Gaps 3;
QY 37 AAACAGAAATGATGGCAAGACTGCTAAGAACATCTTTGCTTTGCTCTTCTTGGCCTC 96
DB 326 AGATAGAAATGACAGCAAGGCTTAAGACAGTTCTTTCTTCTTCTTCTTCTTCTTCT 385
QY 97 TTTGGGTGCTGGGGCAGCAACAATTTTCATGCAAAATGAAGAGGGAAGCTGTGGAC 156
DB 386 TCTGGGTCTCTGGGACACCAAGAAATCTCATGAGAAATGAATATGTTGGAAGCTGTGGAC 445
QY 157 TGGTTTACTTTTATAGTTACTTAAAGACAAACAGAAAGTGGAGAGACTGGTTA 216
DB 446 TGGTTTACTTTTATAGTTACTTAAAGACAAACAGAAAGTGGAGAGACTGGTTA 505
QY 217 GAGTACCTGTACTAGACTCTTACAATCTAGAACTGGAGGAAGTGGAGCAACTAATGAAT 276
DB 506 CAGTACCTGTACTAGACTCTTACAATCTAGAACTGGAGGAAGTGGAGCAACTAATGAAT 565
QY 277 GACACAGAGTGTGGGAGGACATTACACAGCTATATGAAGCATATGCTCTTAAG 336
DB 566 AGCACCAGGAGTGCTCTGGGGAAGACCTTACAGCATCTGTATGACACATTAATCCAG 625
QY 337 ACTAACACAGAGCTTCTTATATATATGATGAGTCCCTTAAACCTGTGAATACAGT 396
DB 626 AATGACA---CAGCCTATCTATATACAGTGTGTCCTTGGATCTGTGATTTACAC 682
QY 397 AGAAATATGACACACCAAGGTTTACTGTGCTGTGGAACAGAGTTCAAGGTTCTGGCTG 456

683 AGACAGTATGACATGCCAAAGGTTCTGCTGTATGAAACAGAACGCGAGGTTCTGGTG 742
QY 457 ATTCAATTCATCCCTCAGTTTCCTCCAAATTCGCGAAGAGCTATGATATCCACCCACA 516
DB 743 ATACACTCTCTTCCCAAGTTTCCCCAGTTT-----ATGGCTATGATGATCCCACTCG 796
QY 517 GCGAGACCAATGACAAAGTGGCATCTGCATCACTTTCAAGTACACCAAGTATGAGCA 576
DB 797 GCGAGGCGATATGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGGAA 856
QY 577 ATAGATTCTCAGCTCTTTGGTCTGCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 636
DB 857 ATAGATTCTCAGCTCTTTGGTCTTACAAACCAACATCTACAGCTCTTCAATCCCAAGCAC 916
QY 637 TTTTACCAGGAGCTCATTACATGCCCCCAGCTGCAACCCAGGCGAGCTCATCAGAGATT 696
DB 917 TTTTCACTGGAAACTTATCTACATGCCCCCGGATGTGCGAACTCCAGTTCTTTAAGATC 976
QY 697 CCTGGCAGGCTCTTCCACCACTTTCAGTTCGCGCCCGCAGGACCAAAATTCCTCCATTTGCA 756
DB 977 CCTGTCGGTACCTCGCTGAACTGCACTCAGCCCGCAGGCTTAAACTTGGTCCATTTTGA 1036
QY 757 AAGTCGATCTTTTCTTGTGACGATCTTTGACGCTGAGTGGCTCAACGCTGAGACACA 816
DB 1037 AAATCAAGTTTATATCTGATGATCTTTTACAGGATGATAGCTCAAAAGTTGAAGACA 1096
QY 817 CACTTGTATACAGAAACCTGGCAGCAAGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 876
DB 1097 CATTGTTAGCACAACCTGGCAGCAAGAAAGACAGAGCTTCTTCAAACTGCTCCCTG 1156
QY 877 CTTTACCAGTGTCTACAATATATAAAGCAATTAATATATACGACATCTTTATTTTCA 936
DB 1157 CTTTACCAGTGTCTACAATATATAAAGCAATTAATATATACGAGTCTTTTACTTCA 1216
QY 937 TATCAAGATACGCGAGTGTGTATTTTCCCAAGGACCAAAATTCGCTGACATGT 996
DB 1217 CGCCAAGACCAATTCAAATGGTGTGTTTCCATAAAGGCTCCGCAATTCGCTGACCTGC 1276
QY 997 ATTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAGTGGAGGATTCATTTGTACC 1056
DB 1277 ATTGGAGACCTTAATCGGAGTCCACCAAGCTTTCAGAGTGGAGGATTCATTTGTACC 1336
QY 1057 CAGAAATTCGCAAAATTTACCAAGCATTTTCAAGGATAGTATTATATATGAAGCTGTAG 1116
DB 1337 AAGATCACTACATTTTACCAGGCAATTTTCAAAATATATATCTCCGTTATGGGTTCTGTA 1396
QY 1117 TAACTTGGTGAAGGACACAGTACTATCATTTGAAACCTTGAATGGGTTCTTCTTCC 1176
DB 1397 TAACTTGGTGAAGGACACACCTTCTCTTCTTGAACACCTGCGCATCGGAACATCTCGC 1456
QY 1177 ATTACACC--TTCTTTATATTTTAAAGGCTGTGAAT 1211
DB 1457 CTTGATCTGTCTCCATTAATTTCAAGGCTTCTGAGT 1493

RESULT 7
AI802154/c
LOCUS
DEFINITION
AI802154
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
676 bp mRNA linear EST 16-DEC-1999
Cx30d07.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2271085.3,
similar to TR:O90759 O90759 DEOXYRIBONUCLEASE II.; mRNA sequence.
AI802154
GI:5367626
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 676)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 982 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 456.
 Location/Qualifiers
 1. .676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2211095"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES

source
 Query Match 49.0%; Score 621.8; DB 9; Length 676;
 Best Local Similarity 97.5%; Pred. No. 2.8e-163;
 Matches 663; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
 574 GCAATAGATCTCAGCTCTGGTGTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 633
 676 GCAATAGATCTCAGCTCTGGTGTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 617
 634 ACCTTTACACAGAGCTCATTCATGCCCCCAGCTGTGCACGGGCGAGCTCATCAGAG 693
 616 ACCTTTACACAGAGCTCATTCATGCCCCCAGCTGTGCACGGGCGAGCTCATCAGAG 557
 694 ATTCTCGCAGGCTCCTCCACACATTCAGTCCGCCCGCAGGACAAAATTCCTCCATTTT 753
 556 ATTCTCGCAGGCTCCTCCACACATTCAGTCCGCCCGCAGGACAAAATTCCTCCATTTT 500
 754 GCAGTCGGATTTCTTTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 813
 499 GCAGTCGGATTTCTTTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 440
 814 ACACACTGTGTTAACAAACCTGGCAGCGAAGAACAGCAAGAGTTCCTTCAAACTGCTCC 873
 439 ACACACTGTGTTAACAAACCTGGCAGCGAAGAACAGCAAGAGTTCCTTCAAACTGCTCC 380
 874 CTTCCTTACCAGTCTACCAATATAAAGCAATTAATATACAGACACTCTTATTTTCAGT 933
 379 CTTCCTTACCAGTCTACCAATATAAAGCAATTAATATACAGACACTCTTATTTTCAGT 320
 934 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAAGGGCACCAAAAATCGCTGGACA 993
 319 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAAGGGCACCAAAAATCGCTGGACA 260
 994 TGTATTGGAGACTTAATCGGAGTCCACACAGCTTCAGAGCTTCAGAGGAGATTCATTGT 1053
 259 TGTATTGGAGACTTAATCGGAGTCCACACAGCTTCAGAGCTTCAGAGGAGATTCATTGT 200
 1054 ACCCAGAAATGGCAATTTACCAAGCATTTCAAGGATTTAGTATTATCTATCAAGACTGT 1113
 199 ACCCAGAAATGGCAATTTACCAAGCATTTCAAGGATTTAGTATTATCTATCAAGACTGT 140
 1114 AAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAAACCTTGACAATGGGTCTTCT 1173

ORIGIN

Query Match 49.0%; Score 621.8; DB 9; Length 676;
 Best Local Similarity 97.5%; Pred. No. 2.8e-163;
 Matches 663; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
 574 GCAATAGATCTCAGCTCTGGTGTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 633
 676 GCAATAGATCTCAGCTCTGGTGTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 617
 634 ACCTTTACACAGAGCTCATTCATGCCCCCAGCTGTGCACGGGCGAGCTCATCAGAG 693
 616 ACCTTTACACAGAGCTCATTCATGCCCCCAGCTGTGCACGGGCGAGCTCATCAGAG 557
 694 ATTCTCGCAGGCTCCTCCACACATTCAGTCCGCCCGCAGGACAAAATTCCTCCATTTT 753
 556 ATTCTCGCAGGCTCCTCCACACATTCAGTCCGCCCGCAGGACAAAATTCCTCCATTTT 500
 754 GCAGTCGGATTTCTTTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 813
 499 GCAGTCGGATTTCTTTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 440
 814 ACACACTGTGTTAACAAACCTGGCAGCGAAGAACAGCAAGAGTTCCTTCAAACTGCTCC 873
 439 ACACACTGTGTTAACAAACCTGGCAGCGAAGAACAGCAAGAGTTCCTTCAAACTGCTCC 380
 874 CTTCCTTACCAGTCTACCAATATAAAGCAATTAATATACAGACACTCTTATTTTCAGT 933
 379 CTTCCTTACCAGTCTACCAATATAAAGCAATTAATATACAGACACTCTTATTTTCAGT 320
 934 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAAGGGCACCAAAAATCGCTGGACA 993
 319 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAAGGGCACCAAAAATCGCTGGACA 260
 994 TGTATTGGAGACTTAATCGGAGTCCACACAGCTTCAGAGCTTCAGAGGAGATTCATTGT 1053
 259 TGTATTGGAGACTTAATCGGAGTCCACACAGCTTCAGAGCTTCAGAGGAGATTCATTGT 200
 1054 ACCCAGAAATGGCAATTTACCAAGCATTTCAAGGATTTAGTATTATCTATCAAGACTGT 1113
 199 ACCCAGAAATGGCAATTTACCAAGCATTTCAAGGATTTAGTATTATCTATCAAGACTGT 140
 1114 AAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAAACCTTGACAATGGGTCTTCT 1173

Db 139 AAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAAACCTTGACAATGGGTCTTCT 80
 QY 1174 TCCATTACACCTCTCTTATATTTTAAAGCCCTGTGAATATATCTTATATACCTGCATATCAC 1233
 Db 79 TCCATTACACCTCTCTTATATTTTAAAGCCCTGTGAATATATCTTATATACCTGCATATCAC 21
 QY 1234 AAAATAAAACATATTTCTCT 1253
 Db 20 AAAATAAAACATTTTCTCT 1
 RESULT 8
 BGI49286 609 bp mRNA linear EST 05-FEB-2001
 LOCUS nad25g11.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3366717 3,
 DEFINITION similar to TR:Q9QZK9 Q9QZK9 DEOXYRIBONUCLEASE DLAD. ;, mRNA
 sequence.
 ACCESSION BGI49286
 VERSION BGI49286.1 GI:12661316
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 609)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 416.
 Location/Qualifiers
 1. .609
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3366717"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."
 ORIGIN
 Query Match 45.8%; Score 580.4; DB 12; Length 609;
 Best Local Similarity 97.4%; Pred. No. 1.2e-151;
 Matches 590; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 650 TCATTTCATGCCCGCAGCTGTGCACAGGCCAGCTCATCAGAGTTCCTGCGAGCTCC 709
 Db 608 TCATTTCATGCTCCAGCTGTGCACAGGCCAGCTCATCAGAGTTCCTGCGAGCTCA 549
 QY 710 TCACCACACTTCAGTCGGCCCGCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATTTCT 769
 Db 548 TCACCACACTTCAGTCGGCCCGCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATTTCT 489

```

770 TTCTTGACGACATCTTTGAGCGCTGGATGGCTCAACGGCTGACGACACACATTTGTAACAG 829
Db      |||||
488 TTCTTGATGACATCTTTGAGCGCTGGATGGATCAACGGCTGACGACACACATTTGTAACAG 429
QY      |||||
830 AAACCTGGCAGCGAGAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGTCT 889
Db      |||||
428 AAACCTGGCAGCGAGAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGTCT 369
QY      |||||
890 ACATATATAAAGCAATTAATATATACGACACATCTTTATTTTCAAGTCTTCTATCAAGATCAG 949
Db      |||||
368 ACATATATAAAGCAATTAATATATACGACACATCTTTATTTTCAAGTCTTCTATCAAGATCAG 309
QY      |||||
950 CCAAGTGGTGTATTTCCCAAGAGGACCAAAATCGCTGACATGTATTTGGAGACCTAA 1009
Db      |||||
308 CCAAGTGGTGTATTTCCCAAGAGGACCAAAATCGCTGACATGTATTTGGAGACCTAA 249
QY      |||||
1010 ATCGGAGTCCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGACCCAGAAATGGCAAA 1069
Db      |||||
248 ATCGGAGTCCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGACCCAGAAATGGCAAA 189
QY      |||||
1070 TTTACCAAGCATTTCAAGGATAGTATATATATATGAAAGCTGTAAAGTAACTTTGGTGAA 1129
Db      |||||
188 TTTACCAAGCATTTCAAGGATAGTATATATATATGAAAGCTGTAAAGTAACTTTGGTGAA 129
QY      |||||
1130 AGGACACAGGTACTATCATTTGAAACCTTGCATAGGCTCTTCCATTTACACCTTCTT 1189
Db      |||||
128 AGGACACAGGTACTATCATTTGAAACCTTGCATAGGCTCTTCCATTTACACCTTCTT 69
QY      |||||
1190 TATATTTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATAT 1249
Db      |||||
68 TATATTTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATAT 9
QY      |||||
1250 CTCTCA 1255
Db      |||||
8 CTCTCA 3

RESULT 9
AI420898/c
LOCUS
DEFINITION
t030906.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2095162 3'
similar to TR:000115 O00115 HYPOTHETICAL HUMAN PROTEIN R31240_2.1;
mRNA sequence.
AI420898 554 bp mRNA linear EST 28-MAR-1999
t030906.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2095162 3'
similar to TR:000115 O00115 HYPOTHETICAL HUMAN PROTEIN R31240_2.1;
mRNA sequence.
AI420898.1 GI:4266829
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 554)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLINL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1310 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 431.
Location/Qualifiers
1. .554
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

```

```

/db_xref="taxon:9606"
/clone="IMAGE:2095162"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/notes="Organ: prostate; Vector: p773D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```

ORIGIN

```

Query Match 43.6%; Score 553; DB 9; Length 554;
Best Local Similarity 99.8%; Pred. No. 5.8e-144;
Matches 553; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 715 ACACCTTCAGTCGGCCAGGACAAAATTCCTCCATTTTGCAAAGTCGGATTCCTTTCTT 774
Db      |||||
554 ACACCTTCAGTCGGCCAGGACAAAATTCCTCCATTTTGCAAAGTCGGATTCCTTTCTT 495
QY 775 GACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAGACACACCTTTTAAACGAAACC 834
Db      |||||
494 GACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAGACACACCTTTTAAACGAAACC 435
QY 835 TGGCAGCGAAAAGACAGACAGCTTCCTTCAAACTGCTCCCTTCCATTCACCATCTTACCAAT 894
Db      |||||
434 TGGCAGCGAAAAGACAGACAGCTTCCTTCAAACTGCTCCCTTCCATTCACCATCTTACCAAT 375
QY 895 ATAAAGCAATTAATATATATATATATATATATATATATATATATATATATATATATAT 954
Db      |||||
374 ATAAAGCAATTAATATATATATATATATATATATATATATATATATATATATATATAT 315
QY 955 TGGTGTATTTCCCAAAAGGGCCACCAAAATCGCTGGACATGTATTTGGAGACCTTAATCGG 1014
Db      |||||
314 TGGTGTATTTCCCAAAAGGGCCACCAAAATCGCTGGACATGTATTTGGAGACCTTAATCGG 255
QY 1015 AGTCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGACCCAGAAATTTGCAAAATTTAC 1074
Db      |||||
254 AGTCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGACCCAGAAATTTGCAAAATTTAC 195
QY 1075 CAAGCATTTCAAGGATTAAGTATATATATATATATATATATATATATATATATATATAT 1134
Db      |||||
194 CAAGCATTTCAAGGATTAAGTATATATATATATATATATATATATATATATATATATATAT 135
QY 1135 ACAGGTACTATCATTTGAAAACCTTGACAAATGGGTCTTCTTCCATTCACCTTCTTTATAT 1194
Db      |||||
134 ACAGGTACTATCATTTGAAAACCTTGACAAATGGGTCTTCTTCCATTCACCTTCTTTATAT 75
QY 1195 TTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATATATATAT 1254
Db      |||||
74 TTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATATATATAT 15
QY 1255 ATGTTTACCATTTA 1268
Db      |||||
14 ATGTTTACCATTTA 1

```

```

RESULT 10
AW295184/c
LOCUS
DEFINITION
UI-H-B12-aib-e-05-0-UI.sl NCI CGAP_Sub4 Homo sapiens cDNA clone
IMAGE:2728785 3', mRNA sequence.
ACCESSION
AW295184
VERSION
AW295184.1 GI:6701820
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


QY 1213 TACTTAACTCGCATATCAAAAATAAAACATATTTCTCT 1253
 Db 41 TACTTAACTCGCATATCAAAAATAAAACATTTCTCT 1

RESULT 13
 LOCUS A1791998 487 bp mRNA linear EST 13-DEC-1999
 DEFINITION OS07C09.y5 NCI CGAP Lu5 Homo sapiens cDNA clone IMAGE:1604656 5', similar to TR:090759 OS0759 DBOXRIBONUCLEASE II ; mRNA sequence.
 ACCESSION A1791998
 VERSION A1791998.1 GI:5339714
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 487)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other ESTs: os07C09.s1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-f@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
 Original clone citation: see original entry for original citation information
 This 5' resequenced clone has no previous 5' data to verify this new read against
 Insert Length: 947 Std Error: 0.00
 Seq primer: -40RP from Gibco
 High quality sequence stop: 469.

FEATURES
 source Location/Qualifiers
 1..487
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1604656"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu5"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 38.1%; Score 482.8; DB 9; Length 487;
 Best Local Similarity 99.4%; Pred. No. 3e-124;
 Matches 484; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 335 AGAGTAACACACAGCGCTATCTAATATACAAATGAGTCCCTAAACCTGTGAATTACA 394
 Db 1 AGAGTAACACACAGCGCTATCTAATATACAAATGAGTCCCTAAACCTGTGAATTACA 60

QY 395 GTAGAAAAGTATGACACACCAAGGTTTACTGCTGGACACAGTTTCAAGGGTTCTGGC 454
 Db 61 GCAGAAAGTATGACACACCAAGGTTTACTGCTGGACACAGTTTCAAGGGTTCTGGC 120

QY 455 TGATTTCATCCATCCCTCAGTTTCCTCCAAATTCGGGAAGAGGCTATGATTATCCACCCA 514
 Db 121 TGATTTCATCCATCCCTCAGTTTCCTCCAAATTCGGGAAGAGGCTATGATTATCCACCCA 180

QY 515 CAGGGAGACGAAATGGAACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACAGTATGAGG 574
 Db 181 CAGGGAGACGAAATGGAACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACAGTATGAGG 240

QY 575 CAATAGATTCTCAGCTCTTGGTTCGCAACCCCAACGCTATAGCTGTCTCCATCCAGCCA 634
 Db 241 CAATAGATTCTCAGCTCTTGGTTCGCAACCCCAACGCTATAGCTGTCTCCATCCAGCCA 300

QY 635 CCTTTCACAGGAGCTCAATTCAGATGCCGAGCTGTCACAGGGCCAGCTCATCAGAGA 694
 Db 301 CCTTTCACAGGAGCTCAATTCAGATGCCGAGCTGTCACAGGGCCAGCTCATCAGAGA 360

QY 695 TTCTGTGCGAGGCTCTCTCACCACACTTCAGTCGGCCAGGACAAAATTTCTCCATTTTG 754
 Db 361 TTCTGTGCGAGGCTCTCTCACCACACTTCAGTCGGCCAGGACAAAATTTCTCCATTTTG 420

QY 755 CAAAGTCGGATTCTTTTCTTGATGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAGA 814
 Db 421 CAAAGTCGGATTCTTTTCTTGATGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAGA 480

QY 815 CACACTT 821
 Db 481 CACACTT 487

RESULT 14
 LOCUS BX280577 482 bp mRNA linear EST 04-MAR-2003
 DEFINITION IMAGE:2271085, mRNA sequence.
 ACCESSION BX280577
 VERSION BX280577.1 GI:28612442
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 482)
 AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.
 Human Unigeneset - RZPD3
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGE:2271085
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No. 972)
<http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972> Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.

FEATURES
 source Location/Qualifiers
 1..482
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2271085"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized

Library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo.

ORIGIN		FEATURES		source	
Query Match	36.9%;	Score 467.8;	DB 13;	Length 482;	
Best Local Similarity	99.4%;	Pred. No. 4.9e-120;			
Matches 480;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;	
QY	335	AGAGTACACACAGCCTATCTAATATACATGATGAGTCCCTAAACCTCTGATATACA	394		
Db	1	AGAGTACACACAGCCTATCTAATATACATGATGAGTCCCTAAACCTCTGATATACA	60		
QY	395	GTAGAAAGTATGGACACACAAAGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGC	454		
Db	61	GCAGAAAGTATGGACACACAAAGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGC	120		
QY	455	TGATTTCATTCCTCCCTCAGTTTCTCCAAATCCGGAAGAGGCTATGATTATCCACCCA	514		
Db	121	TGATTTCATTCCTCCCTCAGTTTCTCCAAATCCGGAAGAGGCTATGATTATCCACCCA	180		
QY	515	CAGGAGACGAAATGGACAAAGTGGCATCTGCATTAACCTTCAAGTACCAACAGTATGAGG	574		
Db	181	CAGGAGACGAAATGGACAAAGTGGCATCTGCATTAACCTTCAAGTACCAACAGTATGAGG	240		
QY	575	CAATAGATTCAGCTCTTGCTGCTGCAACCCCAAGCTCTATAGTGTCTCCATCCAGCCA	634		
Db	241	CAATAGATTCAGCTCTTGCTGCTGCAACCCCAAGCTCTATAGTGTCTCCATCCAGCCA	300		
QY	635	CTTTTACCAGAGCTCAATTCATGCTCCAGCTGTGCACAGGCTCAGTCTCATCAGAGA	694		
Db	301	CTTTTACCAGAGCTCAATTCATGCTCCAGCTGTGCACAGGCTCAGTCTCATCAGAGA	360		
QY	695	TTCTGGCAGGCTCTCTACACACATTCAGTGGCCCGCAGGACAAATTCCTCCATTTTG	754		
Db	361	TTCTGGCAGGCTCTCTACACACATTCAGTGGCCCGCAGGACAAATTCCT-CAITTTG	419		
QY	755	CAAGTCGGATTTCTTTCTTGACGACATCTTTGACGCTGTGATGCTCAACGGCTGAAGA	814		
Db	420	CAAGTCGGATTTCTTTCTTGACGACATCTTTGACGCTGTGATGCTCAACGGCTGAAGA	479		
QY	815	CAC 817			
Db	480	CAC 482			

RESULT 15
AA988125/c
LOCUS
DEFINITION
os07c09.s1 NCI CGAP Lu5 Homo sapiens CDNA clone IMAGE:1604656 3', similar to TR:000115 000115 HYPOTHETICAL HUMAN PROTEIN R31240_2. ; mRNA sequence.
AA988125
AA988125.1 GI:3173489
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 446)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/hbrp/image/image.html
Insert length: 947 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 365.

FEATURES		Location/Qualifiers	
1.	446	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:1604656"	
		/tissue_type="carcinoid"	
		/lab_host="DH10B"	
		/clone_lib="NCI-CCGAP Lu5"	
		/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from a neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."	

ORIGIN		Query Match		Score 441.8;		DB 9;		Length 446;	
Best Local Similarity		99.3%;		Pred. No. 9.6e-113;					
Matches 443;		Conservative 0;		Mismatches 3;		Indels 0;		Gaps 0;	
QY	808	CTGAAGACACACTTGTAAACAGAAACCTGGCAGCGAAGAAAGACAGAGCTTCTTCAAAAC	867						
Db	446	CTGAAGACACACTTGTAAACAGAAACCTGGCAGCGAAGAAAGACAGAGCTTCTTCAAAAC	387						
QY	868	TGCTCCCTTCTTACCATGTCTACAATAAAGCAATTAATATACGACACTCTTAT	927						
Db	386	TGCTCCCTTCTTACCATGTCTACAATAAAGCAATTAATATACGACACTCTTAT	327						
QY	928	TTTCAGTCTTATCAAGATCAGCCAGTGGTGTATTTCCCAAGGGGACCAAAATGCG	987						
Db	326	TTTCAGTCTTATCAAGATCAGCCAGTGGTGTATTTCCCAAGGGGACCAAAATGCG	267						
QY	988	TGGACATGTATTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAGAAAGTGGAGATTC	1047						
Db	266	TGGACATGTATTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAGAAAGTGGAGATTC	207						
QY	1048	ATTTGTACCCAGAAATGGCAATTTTACCAAGCATTTCAAGGATTAGTATTACTATGAA	1107						
Db	206	ATTTGTACCCAGAAATGGCAATTTTACCAAGCATTTCAAGGATTAGTATTACTATGAA	147						
QY	1108	AGCTGTAAGTAACTTGTGAAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGGG	1167						
Db	146	AGCTGTAAGTAACTTGTGAAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGGG	87						
QY	1168	TCTTCTTCATTCACCTTCTTATTTTAAAGCGCTGTGAATATATCTTATAACTGCA	1227						
Db	86	TCTTCTTCATTCACCTTCTTATTTTAAAGCGCTGTGAATATATCTTATAACTGCA	27						
QY	1228	TATCAAAAATAAATATATTTCTCT	1253						
Db	26	TATCAAAAATAAATATATTTCTCT	1						

Search completed: October 14, 2004, 05:24:16
Job time : 3099.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:11:29 ; Search time 71.8017 Seconds
(without alignments)
1404.834 Million cell updates/sec

Title: US-10-790-589-4
Perfect score: 1933
Sequence: 1 MMARLLRTSFALLFLGLFGV.....QNWQIQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1933	100.0	357	4	AAE13015 Human deo
2	1933	100.0	357	5	ABG61821 Prostate
3	1926	99.6	361	4	AAE13015 Human deo
4	1293.5	66.9	354	4	AAE13015 Human deo
5	1293.5	66.9	354	4	AAE13015 Human deo
6	613	31.7	360	2	AAW31495 Human DNA
7	613	31.7	360	2	AAW31495 Human DNA
8	459	23.7	276	2	AAW37921 Bos tauru
9	449	23.2	366	4	ABG63496 Drosophil
10	229	11.8	192	6	ABG72082 Mouse end
11	98.5	5.1	391	6	ABU41838 Protein e
12	98.5	5.1	702	2	ABU41838 Protein e
13	98.5	5.1	1272	4	ABG03108 Chlamydia
14	97.5	5.0	697	5	ABG91036 Chlamydia
15	97.5	5.0	697	7	ADDA43730 Chlamydia
16	97	5.0	426	2	AAE13015 Human deo
17	96.5	5.0	397	4	ABG11583 Raphanus
18	96.5	5.0	661	2	ABG11583 Raphanus
19	95.5	4.9	308	5	ABP29969 Streptoco
20	95.5	4.9	328	5	ABP27445 Streptoco
21	95	4.9	595	2	AAE13015 Human deo
22	95	4.9	595	2	AAE13015 Human deo
23	95	4.9	595	2	AAE13015 Human deo
24	95	4.9	595	2	AAE13015 Human deo
25	95	4.9	687	5	AAE13015 Human deo

26	95	4.9	704	4	AAE13015 standard; protein; 357 AA.
27	95	4.9	704	5	AAE13015; (first entry)
28	95	4.9	704	5	Human deoxyribonuclease (DNase) II beta protein.
29	95	4.9	704	5	Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
30	95	4.9	714	5	cystic fibrosis; pulmonary; chromosome 1p22.
31	95	4.9	970	6	Homo sapiens.
32	94.5	4.9	2295	3	WO200175082-A1.
33	93	4.8	526	6	11-OCT-2001.
34	93	4.8	2472	5	02-APR-2001; 2001WO-US010635.
35	93	4.8	6576	6	03-APR-2000; 2000US-00541840.
36	93	4.8	6669	6	19-MAY-2000; 2000US-00574942.
37	93	4.8	6669	7	(DART-) DARTMOUTH COLLEGE.
38	93	4.8	6700	6	Eastman AR, Krieser RJ;
39	92.5	4.8	638	5	WPI; 2001-662972/76.
40	92	4.8	517	4	N-PSDB; AAD21289.
41	92	4.8	524	4	New CDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
42	92	4.8	641	7	DNA present in the mucous plugs in the lungs of cystic fibrosis patients.
43	92	4.8	641	7	Claim 5; Page 16-17; 21pp; English.
44	91	4.7	704	5	The invention relates to deoxyribonuclease (DNase) II beta proteins and
45	91	4.7	1035	4	their corresponding cDNAs. The DNase II beta may be useful to digest DNA

ALIGNMENTS

RESULT 1					
AAE13015					
ID	AAE13015	standard; protein; 357 AA.			
XX	AC	AAE13015;			
XX	AC	AAE13015;			
DT	28-JAN-2002	(first entry)			
XX	XX	Human deoxyribonuclease (DNase) II beta protein.			
DE	Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;				
KW	cystic fibrosis; pulmonary; chromosome 1p22.				
XX	XX	Homo sapiens.			
OS	Homo sapiens.				
XX	XX	WO200175082-A1.			
PN	WO200175082-A1.				
XX	XX	11-OCT-2001.			
PD	XX	11-OCT-2001.			
PF	XX	02-APR-2001; 2001WO-US010635.			
XX	XX	03-APR-2000; 2000US-00541840.			
PR	XX	19-MAY-2000; 2000US-00574942.			
XX	XX	(DART-) DARTMOUTH COLLEGE.			
XX	XX	Eastman AR, Krieser RJ;			
PI	XX	WPI; 2001-662972/76.			
XX	XX	N-PSDB; AAD21289.			
DR	XX	New CDNA encoding a deoxyribonuclease II beta enzyme useful for degrading			
XX	XX	DNA present in the mucous plugs in the lungs of cystic fibrosis patients.			
PT	XX	Claim 5; Page 16-17; 21pp; English.			
PS	XX	The invention relates to deoxyribonuclease (DNase) II beta proteins and			
XX	XX	their corresponding cDNAs. The DNase II beta may be useful to digest DNA			
CC	XX	in the mucous plugs in lungs of cystic fibrosis patients and so reduce			
CC	XX	their viscosity. The present sequence is human DNase II beta protein. The			
CC	XX	human DNase II beta gene is located at chromosome 1p22			
XX	XX	Sequence 357 AA;			
SQ	XX	Query Match 100.0%; Score 1933; DB 4; Length 357;			
		Best Local Similarity 100.0%; Pred. No. 9, 1e-174;			
		Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQKESGETGLEYL 60
 DB 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQKESGETGLEYL 60
 QY 61 YLDSTTSRWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVKPKVNYSRKY 120
 DB 61 YLDSTTSRWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVKPKVNYSRKY 120
 QY 121 GHTKGLLLMNRVQGFLLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180
 DB 121 GHTKGLLLMNRVQGFLLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180
 QY 181 QLLVCNPNVYSCSIPATFHQELIHPOLCTRASSSEIPGRLLITLQSAQOGKFLHFAKSD 240
 DB 181 QLLVCNPNVYSCSIPATFHQELIHPOLCTRASSSEIPGRLLITLQSAQOGKFLHFAKSD 240
 QY 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
 DB 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
 QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGFICTQNWQIYQAFQGLVLYESCK 357
 DB 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 2

ABG61821
 ID ABG61821 standard; protein; 357 AA.

XX AC ABG61821;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #22.

XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX PN WO200230269-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX PR 13-OCT-2000; 2000US-00687576.

XX PR 08-DEC-2000; 2000US-00733288.

XX PR 08-DEC-2000; 2000US-00733742.

XX PR 24-JAN-2001; 2001US-0263957P.

XX PR 16-MAR-2001; 2001US-0276791P.

XX PR 16-MAR-2001; 2001US-0276888P.

XX PR 06-APR-2001; 2001US-0281922P.

XX PR 24-APR-2001; 2001US-0286214P.

XX PR 30-APR-2001; 2001US-00847046.

XX PR 04-MAY-2001; 2001US-0288589P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX DR WPI; 2002-471335/50.

XX DR N-PSDB; ABK92136.

XX PT Detecting a prostate cancer-associated transcript in a cell in a patient,
 XX PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 XX PT by determining if prostate cancer-associated genes are expressed in a
 XX PT prostate tissue.
 XX PS Claim 27; Page 317; 436pp; English.
 XX CC The present invention relates to methods of detecting a prostate cancer-
 XX CC associated transcript in a cell from a patient. The method comprises
 XX CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
 CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins
 XX SQ Sequence 357 AA;

Query Match 100.0%; Score 1933; DB 5; Length 357;

Best Local Similarity 100.0%; Pred. No. 9.1e-174; Indels 0; Gaps 0;
 Matches 357; Conservative 0; Mismatches 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQKESGETGLEYL 60

DB 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQKESGETGLEYL 60

QY 61 YLDSTTSRWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVKPKVNYSRKY 120

DB 61 YLDSTTSRWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVKPKVNYSRKY 120

QY 121 GHTKGLLLMNRVQGFLLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180

DB 121 GHTKGLLLMNRVQGFLLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180

QY 181 QLLVCNPNVYSCSIPATFHQELIHPOLCTRASSSEIPGRLLITLQSAQOGKFLHFAKSD 240

DB 181 QLLVCNPNVYSCSIPATFHQELIHPOLCTRASSSEIPGRLLITLQSAQOGKFLHFAKSD 240

QY 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300

DB 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

DB 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 3

ABG72417
 ID AAB72417 standard; protein; 361 AA.

XX AC AAB72417;

XX DT 03-MAY-2001 (first entry)

XX DE Human DNase.

XX KW Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
 XX KW infectious disease.

XX OS Homo sapiens.

XX PN WO200112793-A1.

XX PD 22-FEB-2001.

XX PF 01-MAY-2000; 2000WO-JP002893.

XX PR 17-AUG-1999; 99JP-00230870.

XX PA (TANU/) TANUMA S.

XX PI Tanuma S, Shikawa D;

XX WPI; 2001-218348/22.

XX DR N-PSDB; AAF60716.

XX Acidic deoxyribonuclease capable of divalent cation-independent cleavage
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX Claim 8; Page 52-53; 61pp; Japanese.

XX The present sequence is a human deoxyribonuclease (DLAD), which is an
 CC endonuclease. DLAD is capable of divalent cation-independent cleavage of
 CC DNA under acidic conditions. This protein can be used as a substitute for
 CC DNase I in treating cystic fibrosis, and is useful in the prevention and
 CC treatment of infectious diseases

XX SQ Sequence 361 AA;

Query Match 99.6%; Score 1926; DB 4; Length 361;

Best Local Similarity 99.7%; Pred. No. 4.2e-173;

Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60

Db 5 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 64

QY 61 YLSDTTSRWKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120

Db 65 YLSDTTSRWKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 124

QY 121 GHTKGLLLNRVQGFVLIHSIPQPPPIPERGYDPPPTGRNGSGICITTKYNOYEIDS 180

Db 125 GHTKGLLLNRVQGFVLIHSIPQPPPIPERGYDPPPTGRNGSGICITTKYNOYEIDS 184

QY 181 QLLVCPNVYSCSIPATFHOELHMPOLCTRASSSEIPGRLLTLTQSAQOKFLHFAKSD 240

Db 185 QLLVCPNVYSCSIPATFHOELHMPOLCTRASSSEIPGRLLTLTQSAQOKFLHFAKSD 244

QY 241 SFLLDIFAAWMAORLKLTLTETWQRQELPNSCSLPYHVYNIKAIKLSRHSYFSSYQD 300

Db 245 SFLLDIFAAWMAORLKLTLTETWQRQELPNSCSLPYHVYNIKAIKLSRHSYFSSYQD 304

QY 301 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357

Db 305 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 361

RESULT 4

AAE13014

ID AAE13014 standard; protein; 354 AA.

AC AAE13014;

XX 03-MAY-2001 (first entry)

DT Murine DNase.

XX Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
 KW infectious disease.

XX Mus musculus.

XX WO200112793-A1.

XX 22-FEB-2001.

XX 01-MAY-2000; 2000WO-JP002893.

XX 17-AUG-1999; 99JP-00230870.

XX (TANU/) TANUMA S.

XX Tanuma S, Shiohawa D;

XX WPI; 2001-218348/22.

XX N-PSDB; AAE60715.

XX

PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX Claim 3; Page 48-49; 61pp; Japanese.

XX The present sequence is a murine deoxyribonuclease (DLAD), which is an
 CC endonuclease. DLAD is capable of divalent cation-independent cleavage of
 CC DNA under acidic conditions. This protein can be used as a substitute for
 CC DNase I in treating cystic fibrosis, and is useful in the prevention and
 CC treatment of infectious diseases

XX SQ Sequence 354 AA;

Query Match 66.9%; Score 1293.5; DB 4; Length 354;

Best Local Similarity 66.4%; Pred. No. 2.4e-113;

Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60

Db 1 MTAKPLRTVLSLLFFALLSGVLGTPEISCRNEYGEAVDWFYKLPKRTSKASEAGLYL 60

QY 61 YLSDTTSRWKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120

Db 61 YLSDTTSRWKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 119

QY 121 GHTKGLLLNRVQGFVLIHSIPQPPPIPERGYDPPPTGRNGSGICITTKYNOYEIDS 180

Db 120 GHAKGLLVNRTQGFVLIHSVPKPPV--HGXEYPTSGRRYQGTGCTTGYSQFEIDF 177

QY 181 QLLVCPNVYSCSIPATFHOELHMPOLCTRASSSEIPGRLLTLTQSAQOKFLHFAKSD 240

Db 178 QLLVCPNVYSCSIPATFHOELHMPOLCTRASSSEIPGRLLTLTQSAQOKFLHFAKSD 237

QY 241 SFLLDIFAAWMAORLKLTLTETWQRQELPNSCSLPYHVYNIKAIKLSRHSYFSSYQD 300

Db 238 FYDDIFTGWIAGKLLHLLAQTWQKKQELPNSCSLPYHVYNIKLSIGVTSKYSFSSRQD 297

QY 301 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357

Db 298 HSKWCYSIKGSANRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 354

RESULT 5

AAE13014

ID AAE13014 standard; protein; 354 AA.

XX AAE13014;

XX 28-JAN-2002 (first entry)

XX Mouse deoxyribonuclease (DNase) II beta protein.

XX Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
 KW cystic fibrosis; pulmonary.

XX Mus sp.

XX WO200175082-A1.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US010635.

XX 03-APR-2000; 2000US-00541840.

XX 19-MAY-2000; 2000US-00574942.

XX (DART-) DARTMOUTH COLLEGE.

XX Eastman AR, Krieser RJ;

XX WPI; 2001-662972/76.

DR N-PSDB; AAD21288.
XX New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.
XX
XX Claim 5; Page 14-15; 21pp; English.
XX
XX The invention relates to deoxyribonuclease (DNase) II beta proteins and
CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA
CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce
CC their viscosity. The present sequence is mouse DNase II beta protein
XX
XX Sequence 354 AA;
XX
Query Match 66.9%; Score 1293.5; DB 4; Length 354;
Best Local Similarity 66.4%; Pred. No. 2.4e-113;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;
QY 1 MWARLLTSFALLFLGLGVGAATISCRNEEGKAVDMFTFYKLPKRONKESGETGLLEYL 60
Db 1 MTKAPLRTVLGLFALSGVLGTPETISCRNRYGEADVFIYKLPKRTSKASEAGLYL 60
QY 61 YLDSTTRSWRSEQLMNDTKSVLGRTLQOLYEAYASKNNNTAYLIYNDGVPKPVNSKY 120
Db 61 YLDSTRTQWNSKSLYLINSTRSALGRTLQHYDTHNS-TNDTAYLIYNDGVPGSVNSRQY 119
QY 121 GHTGKLLWNRVQGFHLHSIPQPIPEEGYDYPPTGRRNGSGICITFKYNOYEADS 180
Db 120 GHAKGLLYWNRVQGFHLHSIPQPIPEEGYDYPPTGRRNGSGICITFKYNOYEADS 177
QY 181 QLLVCPNPNVYSCITPATFOELIHPQLCTRASSSEIPGRLLITLQSAQGGKFLHFAKSD 240
Db 178 QLLVCPNPNVYSCITPATFOELIHPQLCTRASSSEIPGRLLITLQSAQGGKFLHFAKSD 237
QY 241 SFLLDDIFAAWMAQRKLTLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYPSYQD 300
Db 238 FYTDDIFGWIQAQKLTLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYPSYQD 297
QY 301 HAKWCISOKGTNRWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 357
Db 298 HSKWCVSIGSARNWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 354

RESULT 6
AAW31495
ID AAW31495 standard; protein; 360 AA.
XX
XX AAW31495;
XX
XX 11-MAY-1998 (first entry)
XX
XX Human DNase II protein.
XX
XX DNase II; human; deoxyribonuclease; pulmonary disease; diagnostic;
XX cellular DNA debris.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..16
XX FT /label= signal_peptide
XX FT Protein 17..360
XX FT /note= "DNase II"
XX FT Modified-site 86..88
XX FT /note= "Asn is N-glycosylated"
XX FT Modified-site 212..214
XX FT /note= "Asn is N-glycosylated"
XX FT Modified-site 266..268
XX FT /note= "Asn is N-glycosylated"
XX FT Modified-site 290..292
XX FT /note= "Asn is N-glycosylated"
XX
XX WO9740134-A2.

XX 30-OCT-1997.
XX
XX 23-APR-1997; 97WO-US0006664.
XX
XX 25-APR-1996; 96US-00639294.
XX
XX (GBTH) GENENTECH INC.
XX
XX Baker KP, Baron WF;
XX
XX WPI; 1997-535820/49.
XX
XX N-PSDB; AAV02903.
XX
XX Human DNase II and related nucleic acids - useful in protein production
XX e.g. for therapeutic use to treat systemic lupus erythematosus and
XX pulmonary diseases e.g. cystic fibrosis.
XX
XX Claim 1; Fig 1; 30pp; English.
XX
XX This sequence represents a novel human deoxyribonuclease, DNase II. This
XX protein is useful to reduce the viscoelasticity of DNA-containing
XX material, e.g. mucus, and used to treat patients with pulmonary diseases
XX or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.
XX It may also improve antibiotic efficacy in the treatment of abscesses,
XX infected lesions etc., provide treatment in non-infected conditions in
XX which there is an accumulation of cellular DNA debris, e.g.
XX pyelonephritis and be used to degrade DNA in biological samples or in
XX or ex vivo gene therapy, and antisense oligonucleotides can be produced
XX from the nucleic acids which can bind to and prevent expression of
XX nucleic acid within cells. The nucleic acids encoding human DNase II
XX used for hybridisation assays for nucleic acids sharing substantial
XX in a sample, or to identify and isolate nucleic acids sharing substantial
XX sequence identity (e.g. encoding naturally-occurring allelic variants of
XX human DNase II). The antibodies can be used to detect and measure human
XX DNase II in tissues or clinical samples, and in the purification of human
XX DNase II. The nucleic acids enable production of human DNase II by
XX recombinant DNA methods in quantities sufficient for clinical use, not
XX previously possible
XX
XX Sequence 360 AA;
XX
Query Match 31.7%; Score 613; DB 2; Length 360;
Best Local Similarity 36.9%; Pred. No. 5e-49;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;
QY 13 LFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPK-RONKESGETGLVLYLDSTTRSWRK 71
Db 4 LLLAALLCVPAGALTCTYGDGQPVDMFVYKLPALRGSGEAAQRLQYKILDESSGWRD 63
QY 72 SEQLMNDTKSVLGRTLQOLYEAYASKNNNTAYLIYNDGVPKP-VNYSRKYGHTKGLLW 129
Db 64 GRALINSEGAVERSLQELYR---SNTSLAFLIYNDQPPQPSKADSSMRGHTKGLVLL 120
QY 130 NRVOGFHLHSIPQF-PTPIPEGYDYPPTGRRNGSGICITFKYNOYEADQLVLCNPN 188
Db 121 DHGGFVLVHSVNFNPPFPAASAAVSWPHSACTYGTQTLTLCVSPFPAQFSGMKQLTYTPW 180
QY 189 VYSCIPATFOELIHPQLCTRASSSEIPGRLLITLQSAQGGKFLHFAKSDFLDIPA 248
Db 181 VYNYQLGIFAQEPFDLENVKGHVSOEPWNSSITLTSQAGAVFQSFAKFKFGDGLYS 240
QY 249 AWAQRLKTLTLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
Db 241 GHLAALGTLNQLQVFWHTKVTGILPNSCSDIMQVLNVNIAPFGPAGPSFNSTEDSHKWCV 300
QY 307 SOKGTNRWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 356
Db 301 SPKGP---WTCVGMNRNQGEGRGGTLCAQLPALWKAFOPLVKNYQPC 347

RESULT 7

AAW37920
ID AAW37920 standard; protein; 360 AA.
XX
AC AAW37920;
XX
DT 11-SEP-1998 (first entry)
XX
DE Homo sapiens DNase II.
XX
KW DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
KW controlled cell death; apoptosis; metamorphosis; cell turnover;
KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
KW autoimmune disease; diagnosis; anticancer agents;
KW enhanced chromosomal rearrangement; chromosome instability.
XX
OS Homo sapiens.
XX
PN WO9816659-A1.
XX
PD 23-APR-1998.
XX
PF 09-OCT-1997; 97WO-US018262.
XX
PR 15-OCT-1996; 96US-0028539P.
XX
PA (DART-) DARTMOUTH COLLEGE.
XX
PI Eastman A, Krieser R;
XX
PI WPI; 1998-251301/22.
DR N-PSDB; AAV29137.
XX
XX Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
PT e.g. to induce apoptosis in tumour cells, identify agents modulating
PT apoptosis and digest DNA in sputum of cystic fibrosis patients.
PT
XX Claim 5; Page 18; 29pp; English.
XX
XX The sequence is that of deoxyribonuclease II (DNase II). The enzyme is
CC useful to digest DNA, e.g. in the lung sputum of cystic fibrosis patients
CC to reduce sputum viscosity. It (or fragments) are also useful to produce
CC antibodies, e.g. to study DNase II expression in cells. DNase II has
CC recently been linked with DNA fragmentation in the early stages of
CC controlled cell death (apoptosis), a process critical to homeostasis
CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead
CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
CC whilst too little can lead to cancer or autoimmune diseases. The
CC antibodies raised against it can therefore be used to diagnose apoptotic
CC stages in selected cells, by contacting cells with the antibody,
CC detecting binding of the antibody with DNase II and determining DNase II
CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic
CC agents, e.g. anticancer agents to promote apoptosis in cells. The
CC antibodies can also be used to identify cells susceptible to premature
CC death. The cDNA encoding it is useful to identify agents modulating
CC apoptosis in cells, by treating cells with an agent, transfecting cells
CC with cDNA and monitoring apoptosis compared with untreated cells.
CC Inhibitors identified may be useful in preventing diseases relating to
CC enhanced chromosomal re-arrangement. Vectors comprising the cDNA can be
CC used to induce apoptosis in selected cells, e.g. tumour cells or cells
CC involved in autoimmune disorders. Antisense oligonucleotides can be
CC administered to cells to inhibit DNase II expression to reduce chromosome
CC instability associated with cancer
XX
XX Sequence 360 AA;
Query Match 31.7%; Score 613; DB 2; Length 360;
Best Local Similarity 36.9%; Pred. No. 5e-49;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;
XX
XX 13 LFLGLGVLGAATISCENEGKAYDFTYKLPK-PQNKESGHTGLLEYLYLSTTSWRK 71
DB 4 LLLALLCPVAGALTCTGDSQGVDFVYVYKLPALRGSGEAAQRLQYKYLDESQGWDR 63

QY 72 SEQLMNDTKSVLGETLQOLYEAYASKNNNTAYLLIYNDGVPKP--VNYSRKYGHTKGLLLW 129
Db 64 GRALINSPGAVGSRSLQLYR--SNTSQAFLFLYNDPQPQSKAQDSSMEGHTKGVLL 120
QY 130 NRQGFLLIHSIQF-PPPEEGYDYPTGRRNGSGICITFKYNOYEAIQSLLVCNPN 188
Db 121 DHGGFWLHVSVNFPPASSAAYSWPHSACTYQTLTLCVSPFAQSKMGKQKLYTTPW 180
QY 189 VYSCISIPATPHQELIHPOLCTRASSSEIPERLLTTLQSAQGGKFLHFAKSDSFLDDIFA 248
Db 181 VYNYQLEGIFAQEPFDLENVVKGHVSOEPWNSITLTLSQAGAVFQSFARFKSGDDLYS 240
QY 249 AWAQRLKTHLLTFTWQKQELPSNCSLPYHYNIKAIKLSRHS--YFSYQDHAKWCI 306
Db 241 GWLAAALGTNLQVFMHTVGLPNSCSDIWLNVNQAIFPPGAPGSPFNSTEDHSKWCY 300
QY 307 SQKGTNRWTCIGDLNRSPHQAFSGGFICTQNWQIYQAFGLVLYYESC 356
Db 301 SPKGP---WTCVGMNENQGEQGGTLCALPALWKAQPLVKNYQPC 347
RESULT 8
AAW37921
ID AAW37921 standard; protein; 276 AA.
XX
AC AAW37921;
XX
DT 11-SEP-1998 (first entry)
XX
DE Bos taurus DNase II.
XX
KW DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
KW controlled cell death; apoptosis; metamorphosis; cell turnover;
KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
KW autoimmune disease; diagnosis; anticancer agents;
KW enhanced chromosomal rearrangement; chromosome instability.
XX
OS Bos taurus.
XX
PN WO9816659-A1.
XX
PD 23-APR-1998.
XX
PF 09-OCT-1997; 97WO-US018262.
XX
PR 15-OCT-1996; 96US-0028539P.
XX
PA (DART-) DARTMOUTH COLLEGE.
XX
PI Eastman A, Krieser R;
XX
PI WPI; 1998-251301/22.
DR N-PSDB; AAV29138.
XX
XX Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
PT e.g. to induce apoptosis in tumour cells, identify agents modulating
PT apoptosis and digest DNA in sputum of cystic fibrosis patients.
XX
XX Claim 5; Page 19; 29pp; English.
XX
XX The sequence is that of deoxyribonuclease II (DNase II). The enzyme is
CC useful to digest DNA, e.g. in the lung sputum of cystic fibrosis patients
CC to reduce sputum viscosity. It (or fragments) are also useful to produce
CC antibodies, e.g. to study DNase II expression in cells. DNase II has
CC recently been linked with DNA fragmentation in the early stages of
CC controlled cell death (apoptosis), a process critical to homeostasis
CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead
CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
CC whilst too little can lead to cancer or autoimmune diseases. The
CC antibodies raised against it can therefore be used to diagnose apoptotic
CC stages in selected cells, by contacting cells with the antibody
CC detecting binding of the antibody with DNase II and determining DNase II
CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic

CC agents, e.g. anticancer agents to promote apoptosis in cells. The
 CC antibodies can also be used to identify cells susceptible to premature
 CC death. The cDNA encoding it is useful to identify agents modulating
 CC apoptosis in cells, by treating cells with an agent, transfecting cells
 CC with cDNA and monitoring apoptosis compared with untreated cells.
 CC Inhibitors identified may be useful in preventing diseases relating to
 CC enhanced chromosomal re-arrangement. Vectors comprising the cDNA can be
 CC used to induce apoptosis in selected cells, e.g. tumour cells or cells
 CC involved in autoimmune disorders. Antisense oligonucleotides can be
 CC administered to cells to inhibit DNase II expression to reduce chromosome
 CC instability associated with cancer
 XX
 SQ Sequence 276 AA;

Query Match 23.7%; Score 459; DB 2; Length 276;
 Best Local Similarity 35.8%; Pred. No. 1.2e-34;
 Matches 95; Conservative 52; Mismatches 108; Indels 10; Gaps 5;
 98 SNNATYLYNDGVKPVNY--SRKYGHKGLLNVRVQGLIHSIPQFP--IPEEGYDY 154
 1 NSQAFVLYNDQPKSSSKSSRGHTKGVLLDQEGFWLIHSVPNPPRASSAAYSW 60
 155 PPTGRRGQSGICITFKYNOYEADISQLLNCNPNYSICIPATHQELIHPQLCTRASS 214
 61 PPGAQKVGTLICVSPFLTQFLDLSKQLTYTPLYVDHRLGDPGQKFPYLEEVVKGHHV 120
 215 SEIPGRLITLQSAQGGKFLHFAKSDSFLDDIFAAWMAQRLKTHLLTETWORKQELPSN 274
 121 RQGPWNSVTLTSKKGATFQGFAGFNFGDLYSGWLAELGSTLQVFWQSSGILPSN 180
 275 CSLPHVHYN--KATKLSRHSYFSSYQDHAKWCISOKGTKNRWTICIGLNRSPhQAFRSG 332
 181 CSGAQHVFDVQTAFPGAGAPAFNATEDHKKVCVTPKGP---WACVGDWNRQREHGG 237
 333 GFICTQNMQIYQAFQGLVLYYESCK 357
 238 GTLCAQ--MLMKAFLPLVKAWBPCE 260

RESULT 9
 ABB63496
 ID ABB63496 standard; protein; 366 AA.

AC ABB63496;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17280.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li FWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL07599.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 17280; 21pp + Sequence Listing; English.
 PS
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 366 AA;

Query Match 23.2%; Score 449; DB 4; Length 366;
 Best Local Similarity 32.6%; Pred. No. 1.6e-33;
 Matches 120; Conservative 68; Mismatches 148; Indels 32; Gaps 15;
 10 FALLFLGLFGLVGA-ATISCTNREGKAVDWFTFYKLPKR-QNKESGE--TGLEYLYLDST 65
 6 FVLLFVFFYQNEAKSKVSCCKDEAGNDVDMWHLKLPKHQHNLDGKDTSLKLYLVTSQ 65
 66 T-RSWRSEQIMNDTKSVLGETLOOLYEAYASKNNNTAYLYNDGVPKPVNYSRYKXGHTK 124
 66 NYDTWQSGKFTISDPLSLPAQTLNPLND---DPSHTLLAAYNDAQPNGTVFS--SGGHAK 120
 125 GLLLNVRVQGLIHSIPQFPPIPEEGYDYPTGRRGQSGICITFKYNOYEADISQLLV 184
 121 GVASDGTATWVHSVPKFTIPD--YSYPTSGEQVAQSMCLCVLKGEDEKVV-GQLIV 177
 185 CN-PNVYSCSIP-ATFHQELIHPQLCTRAS---SSIPGRLITLQSAQGGKFLHFAK 238
 178 YNEPHFYQRPNPLATRSDEL--FPSL-ERALHGQWRTESSPKQLEVRSLDGKKFLRGK 234
 239 SDSFLDDIFAAMQAQRLKTHLLTETWORKQELPSNCSLPYHVNINIKALKSRHSY-PSS 297
 235 SGRANVELYADWVAPTLDVSLFVEAWRDGAGNLPNSCDKSKVLNVESISNPELSVDFKT 294
 298 YODHAKWCISQ-----KGTKNRWTICIGLNRSPhQAFRSGGFICTONWQIYQAFQ 348
 295 TQDHSKWAVSRTPTGILIYHVRVGGDWICVGDINQEQGLHGGTGTCHKSARVSNLYRQ 354
 349 LVLYYESC 356
 355 LVTVYDKC 362

RESULT 10
 ABB72082
 ID ABB72082 standard; protein; 192 AA.

AC ABB72082;

DT 10-FEB-2003 (first entry)

DE Mouse endonuclease implicated in switch recombination (Endo-SR).

XX Mouse; enzyme; endonuclease implicated in switch recombination; Endo-SR;
 KW G-rich region; DNA rearrangement; Genetic recombination; B lymphocyte;
 KW Ig constant region; immune system; B cell; ligand specificity;
 KW gene therapy; tumour; cancer; leukaemia; lymphoma; apoptosis;
 KW programmed cell death; peptidomimetic.

OS Mus sp.

XX Location/Qualifiers
 FH Key
 FT Misc-difference 142..143
 FT /note= "Encoded by CCTACACCTATCCCTTGTCTATGACCAACAGCTGGA
 FT AGGCTTCTTGCTCAGAAATACCTGAC"

XX US6455250-B1.

XX PD 24-SEP-2002.
 XX PF 11-DEC-1998; 98US-00210422.
 XX PR 11-DEC-1997; 97US-0069205P.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PI Aguilera RJ, Lyon CJ;
 XX DR WPI; 2003-066229/06.
 XX DR N-PSDB; ABS58030.
 XX PT New nucleic acid and its encoded murine endonuclease-SR protein, useful
 XX PT for treating tumors or cancers, e.g. leukemias or lymphomas, or for
 XX PT designing and isolating peptidomimetics or inhibitors of Endo-SR for
 XX PT treating these diseases.
 XX PS Example 5; Col 82; 66pp; English.
 XX CC The invention discloses an isolated nucleic acid molecule, which encodes
 XX CC a murine endonuclease implicated in switch recombination (Endo-SR)
 XX CC protein. The enzyme is useful at cleaving DNA at specific G-rich regions
 XX CC which are implicated in modulating DNA rearrangements. Also disclosed are
 XX CC methods for repairing DNA and modulating genetic recombination in a cell.
 XX CC One example of genetic rearrangement is in mature B lymphocytes which can
 XX CC alter their Ig constant region for another, significantly enhancing the
 XX CC versatility of the immune system by allowing B cells to alter their
 XX CC function without altering their ligand specificity. The endo-SR nucleic
 XX CC acid and protein are useful (using gene therapy) for treating tumours or
 XX CC cancers, e.g. leukemias or lymphomas, as well as the modulation of
 XX CC apoptosis and programmed cell death events. The endo-SR nucleic acid and
 XX CC protein are also useful for designing and isolating peptidomimetics and
 XX CC inhibitors of Endo-SR, which may be employed for treating the diseases.
 XX CC The sequence presented is the mouse endo-SR protein
 XX SQ Sequence 192 AA;
 Query Match 11.8%; Score 229; DB 6; Length 192;
 Best Local Similarity 26.5%; Pred. No. 4e-13;
 Matches 67; Conservative 25; Mismatches 95; Indels 66; Gaps 6;
 QY 6 LRTSFALLFLGVLGALATISCRNEGKAVDWFYFKLPKQKESGETGLVYLDST 65
 Db 1 MATLSILLAALLWV-PAEALSCVGSQGVDFVYVKLPAHSGSRDTPKGLTCKTMDQN 59
 QY 66 TRSWRKSEQLMNDTKSVLGRTLQOLYEAVASKNNNTAYLLYNDGVPK--PWNYSRKYGHT 123
 Db 60 SDGWQDGVGYINSPEGAVGRSLQPLXR---KNSSQLAFLLYNDQPPKSSSTRDSTGHGHT 116
 QY 124 KGLLLNWRVQGFVLIHSIQPFPIPEGVDYPTGRRNGSGICIFKYNQYRAIDSQLL 193
 Db 117 KG-----KQLT 122
 QY 184 VCNPNVYSCSIPATFHEQLHMPQLCTRASSSEI---PGRLLTLOSAGOKFLHAKSD 240
 Db 123 YTVPLVYDKLGEFFAQLK---PDLETVIKQVHLPEWPNSSVILLISQAGATQSPAKFG 179
 QY 241 SFLLDDIFAAWMAQ 253
 Db 180 KFGDDLYSGWLAE 192
 RESULT 11
 ABU41838
 ID ABU41838 standard; protein; 391 AA.
 XX AC ABU41838;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #27365.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Pseudomonas syringae.
 XX XN WO200277183-A2.
 XX XN 03-OCT-2002.
 XX XF 21-MAR-2002; 2002WO-US009107.
 XX XN 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX XN (ELIT-) ELITRA PHARM INC.
 XX XN Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX XN Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PSDB; ACA45708.
 XX CC New antisense nucleic acids, useful for identifying proteins or screening
 XX CC for homologous nucleic acids required for cellular proliferation to
 XX CC isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 69762; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 XX CC the 6213 antisense sequences given in the specification where expression
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX CC encoding a polypeptide whose expression is inhibited by the antisense
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX CC polypeptide or its fragment whose expression is inhibited by the
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX CC proliferation; (7) identifying a gene in an operon required for
 XX CC proliferation; (8) identifying a compound that influences the activity of
 XX CC the gene product or that has an activity against a biological pathway
 XX CC required for proliferation, or that inhibits cellular proliferation; (8)
 XX CC identifying a gene required for cellular proliferation or the biological
 XX CC pathway in which a proliferation-required gene or its gene product lies
 XX CC or a gene on which the test compound that inhibits proliferation of an
 XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX CC product is overexpressed or underexpressed; (12) determining the extent
 XX CC to which each of the strains is present in a culture or collection of
 XX CC strains; or (13) identifying the target of a compound that inhibits the
 XX CC proliferation of an organism. The antisense nucleic acids are useful for
 XX CC identifying proteins or screening for homologous nucleic acids required
 XX CC for cellular proliferation to isolate candidate molecules for rational
 XX CC drug discovery programs, or for screening homologous nucleic acids
 XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 XX CC the target prokaryotic essential genes. Note: The sequence data for this
 XX CC patent did not form part of the printed specification, but was obtained
 XX CC in electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 391 AA;

Query Match 5.1%; Score 98.5; DB 6; Length 391;
 Best Local Similarity 24.0%; Pred. No. 2.3;
 Matches 63; Conservative 33; Mismatches 110; Indels 57; Gaps 11;
 QY 15 LGLFVGLGAATISCRNEGKAVDWFYFKLPKQKESGET-----GLEYLYL-DS 64
 Db 10 LGL-GAMGAATVYQLAKAGVDVIGVDYRAPPHTQSSSHDTRITLSVGEQPLPLVRS 58

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
xx Sequence 1272 AA;

[illegible]

RESULT 14	
ABG91036	
ID	ABG91036 standard; protein; 697 AA.
XX	
AC	ABG91036;
XX	
DT	29-NOV-2002 (first entry)
XX	
DE	Chlamydia trachomatis hypothetical protein #4.
XX	
KW	Gram-negative bacterial bleb; PorB; outer membrane protein;
KW	Chlamydia trachomatis infection; Chlamydia pneumoniae infection;
KW	protective antigen; antibacterial; vaccine.
XX	
OS	Chlamydia trachomatis.
XX	
FPN	WO200262380-A2.
XX	
PPD	15-AUG-2002.
XX	
PPF	08-FEB-2002; 2002WO-EP001356.
XX	
PPR	08-FEB-2001; 2001GB-00003169.
XX	
PPA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;
XX	
DR	WPI; 2002-657510/70.
XX	
DR	N-PSDB; ABS67357.
XX	
PPT	Novel gram-negative bacterial bleb presenting on its surface PorB outer
PPT	membrane protein from Chlamydia trachomatis or protective antigen from
PPT	Chlamydia pneumoniae, useful for preventing Chlamydia infection.
XX	

PS Disclosure; Page 20; 75pp; English.

xx

CC The present invention relates to a new gram-negative bacterial bleb

CC presenting on its surface the PorB outer membrane protein from Chlamydia

CC trachomatis, or a protective antigen from C. pneumoniae. The invention is

CC useful for preventing C. trachomatis or C. pneumoniae infection in a

CC host. The present amino acid sequence represents a Chlamydia trachomatis

CC protein of the invention

CC xx

CC xx

CC Sequence 697 AA:

SQ

Query Match	5.0%;	Score 97.5;	DB 5;	Length 697;
Best Local Similarity	24.9%;	Pred. No. 6.6;		
Matches	62;	Conservative	25;	Mismatches 109; Indels 53; Gaps 12
QY	114	VNYSRK-----YCHTKGLLWN-RVQGF-WLIH-----SIPQFPPIPEGVDYPPPTG	158	
Db	85	VNVRGKILICDYLEYYEDTDCLLTNGRCSLYPWFIGGSTITISPFSSIIHKGYIISTSEG	144	
QY	159	RRNQSGICITCFKNOYEADISQLLCNPNVYSCISIPATFQHEL-----IHMPOLCWRA	212	
Db	145	----PQKHCISGDKYKYS--DSVLSMGPSRLSICNTPVLLPQISIMPMEIPKPPITFRG	200	
QY	213	SSSEIPGLRLITTLQSAQGFLEFAKSDSFLDDTFAAWMAQRLKTHLLTETWQRKQSLP	272	
Db	201	GSQGFGLSYLGVSYSPISKK---HCSTTLFLDGFKKHGIG-----LGYNRRFSSQBNP	250	
QY	273	SNCSLPVHVYNIKAIKLSRHSYFSSYODHAKWCISQKTKNRTWCIGDLNRSPhQAFSG	332	
Db	251	SN-----ALNIK-----SYAH-RLAIDSSGAKDRYELHGDFFSKERAHLAG	292	
QY	333	GFICTQNWQ 341		
Db	293	EFLHDSWE 301		

RESULT 15	
ADD43730	
ID	ADD43730 standard; protein; 697 AA.
XX	
AC	ADD43730;
XX	
XX	
XX	15-JAN-2004 (first entry)
XX	
XX	Chlamydia trachomatis immunogenic protein, SEQ ID No 25.
DE	
DE	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW	gene therapy; antibacterial.
XX	
XX	Chlamydia trachomatis.
OS	
XX	
PN	WO2003049762-A2.
XX	
PD	19-JUN-2003.
XX	
PF	12-DEC-2002; 2002WO-IB005761.
XX	
PR	12-DEC-2001; 2001GB-00029732.
PR	06-AUG-2002; 2002GB-00018233.
PR	14-AUG-2002; 2002GB-00018924.
XX	
XX	(CHIR-) CHIRON SPA.
XX	
PI	Grandi G, Ratti G;
XX	
XX	WPI; 2003-532882/50.
DR	N-PSDB; ADD43731.
DR	
XX	
XX	New immunogenic composition having a protein or encoding nucleic
PPT	acid useful for diagnosing, preventing and/or treating Chlamydia trach
PPT	infection.
PPT	
XX	
PS	Claim 6; SEQ ID NO 25; 164pp; English.

us-10-790-589-4.rag

The invention relates to a novel immunogenic composition comprising a protein or nucleic acid, and an adjuvant, where the protein or nucleic acid comprises any of 131 fully defined amino acid or nucleotide sequences given in the specification, or has 50% or greater sequence identity to it, or their fragments. The protein and/or nucleic acid of the immunogenic composition is useful in the manufacture of a medicament for the treatment or prevention of infection due to Chlamydia trachomatis. The infection is treated or prevented by the medicament eliciting an immune response which is specific to a C. trachomatis elementary body or for neutralising C. trachomatis elementary bodies, hence the immunogenic composition can be used in creating a vaccine. The immunogenic compositions can also be used for the diagnosis of C. trachomatis infection. The nucleic acids of the immunogenic compositions can be used to treat disorders by gene therapy. The immunogenic compositions have antibacterial activity. This sequence represents one of the 131 C. trachomatis proteins with immunogenic properties of the invention.

Sequence 697 AA;

Query Match	5.0%;	Score	97.5;	DB	7;	Length	697;
Best Local Similarity	24.9%;	Pred. No.	6.6;				
Matches	62;	Conservative	25;	Mismatches	109;	Indels	53;
							Gaps
							12;
QY	114	VVYSRK-----	YGHTRGLLIWN-RVQGF-WLIH-----	SIPQFPPIEEGYDYPTTG	158		
Db	85	VVYRGKILICDYLEVEYEDTSCLLTNGKRCSLYPWFICGGTITITSPSSIIHKGYISTSEG	144				
QY	159	RNQSOGICITFKKNOYEADISOLLVCNPNVYSCSIPATHQEL-----	IHMPLQCTRA	212			
Db	145	---PQKHICLSDGILKYSS-DSVLUSMGPSRLICNTVLLLPQLSIMPMLPPPTFRG	200				
QY	213	SSSEIPGELLTLQSAQCKPHEFAKSDSFLLDIFAAWMAQRKTHLLTTLTTWORKQELP	272				
Db	201	GSGFGFLGSLGVSYSPISK--HCSNTLFLDGFKEHGIG-----	LYNMRFSSQENP	250			
QY	273	SNCSLPYHVYNIKAILSRHSYFSSYODHAKWCISQKTKNRWTCIGDLNRSRHOAFPSG	332				
Db	251	SN-----AINIK-----	SYAH-FLAIDSAGKDRYPLHGDFTFSKERAHLAG	292			
QY	333	GFICTQNWQ	341				
Db	293	EFHLSDSWE	301				

Search completed: October 14, 2004, 00:28:08
Job time : 73.8017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:24:00 ; Search time 31.6329 Seconds
(without alignments)
582.636 Million cell updates/sec

Title: US-10-790-589-4

Perfect score: 1933

Sequence: 1 MWARLRTSFALLFLGLFGV.....QNWIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2.6/prodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2.6/prodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2.6/prodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2.6/prodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2.6/prodata/2/iaa/6C COMB.pcp.*
- 6: /cgn2.6/prodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1933	100.0	357	4	US-09-574-942-4
2	1926	99.6	361	4	US-09-807-784B-3
3	1293.5	66.9	354	4	US-09-574-942-2
4	1293.5	66.9	354	4	US-09-807-784B-1
5	613	31.7	360	3	US-09-147-915-3
6	613	31.7	360	3	US-08-639-294-2
7	613	31.7	360	4	US-09-861-034B-2
8	458	23.7	275	3	US-09-147-915-4
9	229	11.8	192	4	US-09-210-422-2
10	96.5	5.0	661	4	US-09-198-452A-36
11	95.5	4.9	426	2	US-08-416-870C-10
12	95.5	4.9	595	1	US-08-468-036-3
13	95	4.9	595	2	US-08-376-843-3
14	95	4.9	704	4	US-09-409-180A-1
15	90.5	4.7	384	4	US-09-071-035-276
16	90.5	4.7	430	4	US-09-071-035-274
17	90.5	4.7	2037	4	US-09-543-681A-5538
18	89.5	4.6	477	4	US-09-134-000C-6123
19	88.5	4.6	542	4	US-09-543-681A-6775
20	88.5	4.6	1456	4	US-09-976-594-168
21	88	4.6	459	1	US-08-220-151-12
22	88	4.6	459	1	US-08-220-151-14
23	88	4.6	459	1	US-08-413-118-12
24	88	4.6	459	1	US-08-413-118-14
25	88	4.6	459	3	US-08-473-446-12
26	88	4.6	459	3	US-08-473-446-14
27	88	4.6	459	3	US-09-213-053-6

28	88	4.6	1394	3	US-09-213-053-2	Sequence 2, Appli
29	84.5	4.4	967	4	US-09-489-039A-8766	Sequence 8766, Ap
30	84.5	4.4	1455	3	US-08-840-062-5	Sequence 5, Appli
31	84	4.3	455	3	US-09-221-235-5	Sequence 5, Appli
32	84	4.3	455	3	US-09-221-928-5	Sequence 5, Appli
33	84	4.3	455	3	US-09-221-527-5	Sequence 5, Appli
34	84	4.3	455	3	US-09-221-236-5	Sequence 5, Appli
35	84	4.3	455	3	US-09-221-416-5	Sequence 5, Appli
36	84	4.3	455	3	US-09-221-245-5	Sequence 5, Appli
37	84	4.3	455	3	US-09-163-115-5	Sequence 5, Appli
38	84	4.3	455	3	US-09-221-528-5	Sequence 5, Appli
39	84	4.3	455	3	US-09-593-553-5	Sequence 5, Appli
40	84	4.3	455	3	US-09-221-237-5	Sequence 5, Appli
41	84	4.3	455	4	US-09-399-588-2	Sequence 5, Appli
42	84	4.3	528	4	US-09-328-352-6385	Sequence 6385, Ap
43	84	4.3	855	2	US-08-482-090-12	Sequence 12, Appli
44	84	4.3	856	2	US-08-481-700B-8	Sequence 8, Appli
45	84	4.3	856	2	US-09-007-383-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-09-574-942-4
; Sequence 4, Application US/09574942
; Patent No. 6358723
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DROXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/574,942
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-574-942-4

Query Match	100.0%;	Score 1933;	DB 4;	Length 357;
Best Local similarity	100.0%;	Pred. No. 1.5e-193;		
Matches 357;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MWARRLRTSFALLFLGLFGVLCGAATISCNERRGKAVDWFYKLPKRNKESGCTGLYL	60	
Db	1	MWARRLRTSFALLFLGLFGVLCGAATISCNERRGKAVDWFYKLPKRNKESGCTGLYL	60	
QY	61	YLDSTTRGWRKSEQLMNDTKSVLGTLOQLYEAYASKNNNTAYLIYNDGVPKPVNSRYK	120	
Db	61	YLDSTTRGWRKSEQLMNDTKSVLGTLOQLYEAYASKNNNTAYLIYNDGVPKPVNSRYK	120	
QY	121	GHTKGLLLNNRVQGFWLHISIPQFPPIPEGDYDPTGRNQSGGICITFKTNOVEAIDS	180	
Db	121	GHTKGLLLNNRVQGFWLHISIPQFPPIPEGDYDPTGRNQSGGICITFKTNOVEAIDS	180	
QY	181	QLLVNPNVYCSIPATFQELIHPOLCTRASSEIPGRLITTTLSAQOGKELHFAKSD	240	
Db	181	QLLVNPNVYCSIPATFQELIHPOLCTRASSEIPGRLITTTLSAQOGKELHFAKSD	240	
QY	241	SFLDDIFAAWMAQRKLTHTLTTWQRKQELPNSCLPVHVYVNIKAIKLSRHSYSSYQD	300	
Db	241	SFLDDIFAAWMAQRKLTHTLTTWQRKQELPNSCLPVHVYVNIKAIKLSRHSYSSYQD	300	
QY	301	HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK	357	
Db	301	HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK	357	

RESULT 2

US-09-807-784B-3
 ; Sequence 3, Application US/09807784B
 ; Patent No. 6653118
 ; GENERAL INFORMATION:
 ; APPLICANT: Tanuma, Sei-ichi
 ; APPLICANT: Shiokawa, Daisuke
 ; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
 ; FILE REFERENCE: 210792
 ; CURRENT APPLICATION NUMBER: US/09/807,784B
 ; CURRENT FILING DATE: 1999-08-17
 ; PRIOR APPLICATION NUMBER: JP 11-230870
 ; PRIOR FILING DATE: 1999-08-17
 ; NUMBER OF SEQ ID NOS: 15
 ; SEQ ID NO 3
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-807-784B-3

Query Match 99.6%; Score 1926; DB 4; Length 361;
 Best Local Similarity 99.7%; Pred. No. 8.4e-193;
 Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MMARLLRTS	FALLFLG	FGVLGA	ATISCR	NEEGK	AVDWF	TFYK	LPK	RK	ESG	ETG	LEYL	60																																				
Db	5	MMARLLRTS	FALLFLG	FGVLGA	ATISCR	NEEGK	AVDWF	TFYK	LPK	RK	ESG	ETG	LEYL	64																																				
Qy	61	YLDSTTR	SWRKSE	QLMNDT	KSVLGR	TQLQ	LYE	AYAS	KSNNT	AYLI	YNDG	VPK	PVNS	SKY 120																																				
Db	65	YLDSTTR	SWRKSE	QLMNDT	KSVLGR	TQLQ	LYE	AYAS	KSNNT	AYLI	YNDG	VPK	PVNS	SKY 124																																				
Qy	121	GHTYGL	LLNNRV	QGF	WLHSI	PP	PE	EGY	DP	PTGR	NG	QSG	GICIT	FKYNOYE	180																																			
Db	125	GHTYGL	LLNNRV	QGF	WLHSI	PP	PE	EGY	DP	PTGR	NG	QSG	GICIT	FKYNOYE	184																																			
Qy	181	QLLV	CNPNV	YSCS	IPAT	FH	QELI	HPOL	CT	RAS	SEI	PGR	LLT	TLQ	SAOQ	OKFLH	FAKSD 240																																	
Db	185	QLLV	CNPNV	YSCS	IPAT	FH	QELI	HPOL	CT	RAS	SEI	PGR	LLT	TLQ	SAOQ	OKFLH	FAKSD 244																																	
Qy	241	SFLDDI	FAA	MAQ	RLK	THLL	TET	W	RK	QEL	P	NSC	SLP	YH	YNI	KA	IKLS	SRHS	YFSS	YQD 300																														
Db	245	SFLG	IFA	MAQ	RLK	THLL	TET	W	RK	QEL	P	NSC	SLP	YH	YNI	KA	IKLS	SRHS	YFSS	YQD 304																														
Qy	301	HAKWC	IS	QK	GT	K	N	R	W	T	C	I	G	D	L	N	R	S	P	H	O	A	F	R	S	G	G	F	I	C	T	Q	N	W	O	I	Y	O	A	F	O	G	L	V	I	Y	E	S	C	K 357
Db	305	HAKWC	IS	QK	GT	K	N	R	W	T	C	I	G	D	L	N	R	S	P	H	O	A	F	R	S	G	G	F	I	C	T	Q	N	W	O	I	Y	O	A	F	O	G	L	V	I	Y	E	S	C	K 361

RESULT 3

US-09-574-942-2
 ; Sequence 2, Application US/09574942
 ; Patent No. 6358723
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan R.
 ; APPLICANT: Krieser, Ronald J.
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
 ; FILE REFERENCE: DC-0137
 ; CURRENT APPLICATION NUMBER: US/09/574,942
 ; CURRENT FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/541,840
 ; PRIOR FILING DATE: 2000-04-03
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 US-09-574-942-2

Query Match 66.9%; Score 1293.5; DB 4; Length 354;
 Best Local Similarity 66.4%; Pred. No. 1.1e-126;
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy	1	MMARLLRTS	FALLFLG	FGVLGA	ATISCR	NEEGK	AVDWF	TFYK	LPK	RK	ESG	ETG	LEYL	60																																				
Db	1	MTAKPLR	TVLSL	FFALSG	VLGP	ET	PEI	S	CR	NE	Y	E	A	V	D	W	F	I	F	YK	LPK	R	K	T	S	K	A	S	E	A	G	L	O	Y	L 60															
Qy	61	YLDSTTR	SWRKSE	QLMNDT	KSVLGR	TQLQ	LYE	AYAS	KSNNT	AYLI	YNDG	VPK	PVNS	SKY 120																																				
Db	61	YLDSTTR	QWNS	LYL	IN	STR	SAL	GRT	L	Q	HL	Y	D	T	H	N	S	-T	N	D	T	A	Y	L	I	N	D	G	V	P	G	S	V	N	S	R	O	Y 119												
Qy	121	GHTYGL	LLNNRV	QGF	WLHSI	PP	PE	EGY	DP	PTGR	NG	QSG	GICIT	FKYNOYE	180																																			
Db	120	GHAKGL	LVN	NR	TQ	GF	WL	HS	VP	K	FP	PV	-H	G	Y	B	T	S	G	R	R	Y	G	O	T	G	I	C	I	T	F	G	S	O	F	E	I	D	F 177											
Qy	181	QLLV	CNPNV	YSCS	IPAT	FH	QELI	HPOL	CT	RAS	SEI	PGR	LLT	TLQ	SAOQ	OKFLH	FAKSD 240																																	
Db	178	QLLV	Q	N	I	Y	S	C	F	I	P	S	F	H	W	K	L	I	Y	M	R	C	A	N	S	S	L	K	I	P	V	R	L	A	E	L	H	S	A	O	G	L	N	F	V	H	F	A	K	S 237
Qy	241	SFLDDI	FAA	MAQ	RLK	THLL	TET	W	RK	QEL	P	NSC	SLP	YH	YNI	KA	IKLS	SRHS	YFSS	YQD 300																														
Db	238	FYTDDI	F	T	G	W	I	A	O	K	L	T	H	L	A	O	T	W	K	K	Q	E	L	P	N	S	C	L	P	Y	H	V	N	I	K	S	I	G	T	S	K	S	F	S	S	R	O	D 297		

RESULT 4

US-09-807-784B-1
 ; Sequence 1, Application US/09807784B
 ; Patent No. 6653118
 ; GENERAL INFORMATION:
 ; APPLICANT: Tanuma, Sei-ichi
 ; APPLICANT: Shiokawa, Daisuke
 ; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
 ; FILE REFERENCE: 210792
 ; CURRENT APPLICATION NUMBER: US/09/807,784B
 ; CURRENT FILING DATE: 1999-08-17
 ; PRIOR APPLICATION NUMBER: JP 11-230870
 ; PRIOR FILING DATE: 1999-08-17
 ; NUMBER OF SEQ ID NOS: 15
 ; SEQ ID NO 1
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-807-784B-1

Query Match 66.9%; Score 1293.5; DB 4; Length 354;
 Best Local Similarity 66.4%; Pred. No. 1.1e-126;
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy	1	MMARLLRTS	FALLFLG	FGVLGA	ATISCR	NEEGK	AVDWF	TFYK	LPK	RK	ESG	ETG	LEYL	60																																				
Db	1	MTAKPLR	TVLSL	FFALSG	VLGP	ET	PEI	S	CR	NE	Y	E	A	V	D	W	F	I	F	YK	LPK	R	K	T	S	K	A	S	E	A	G	L	O	Y	L 60															
Qy	61	YLDSTTR	SWRKSE	QLMNDT	KSVLGR	TQLQ	LYE	AYAS	KSNNT	AYLI	YNDG	VPK	PVNS	SKY 120																																				
Db	61	YLDSTTR	QWNS	LYL	IN	STR	SAL	GRT	L	Q	HL	Y	D	T	H	N	S	-T	N	D	T	A	Y	L	I	N	D	G	V	P	G	S	V	N	S	R	O	Y 119												
Qy	121	GHTYGL	LLNNRV	QGF	WLHSI	PP	PE	EGY	DP	PTGR	NG	QSG	GICIT	FKYNOYE	180																																			
Db	120	GHAKGL	LVN	NR	TQ	GF	WL	HS	VP	K	FP	PV	-H	G	Y	B	T	S	G	R	R	Y	G	O	T	G	I	C	I	T	F	G	S	O	F	E	I	D	F 177											
Qy	181	QLLV	CNPNV	YSCS	IPAT	FH	QELI	HPOL	CT	RAS	SEI	PGR	LLT	TLQ	SAOQ	OKFLH	FAKSD 240																																	
Db	178	QLLV	Q	N	I	Y	S	C	F	I	P	S	F	H	W	K	L	I	Y	M	R	C	A	N	S	S	L	K	I	P	V	R	L	A	E	L	H	S	A	O	G	L	N	F	V	H	F	A	K	S 237
Qy	241	SFLDDI	FAA	MAQ	RLK	THLL	TET	W	RK	QEL	P	NSC	SLP	YH	YNI	KA	IKLS	SRHS	YFSS	YQD 300																														
Db	238	FYTDDI	F	T	G	W	I	A	O	K	L	T	H	L	A	O	T	W	K	K	Q	E	L	P	N	S	C	L	P	Y	H	V	N	I	K	S	I	G	T	S	K	S	F	S	S	R	O	D 297		

Query Match 66.9%; Score 1293.5; DB 4; Length 354;
 Best Local Similarity 66.4%; Pred. No. 1.1e-126;
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

RESULT 5

US-09-147-915-3
; Sequence 3, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krieser, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-147-915-3

Query Match 31.7%; Score 613; DB 3; Length 360;

Best Local Similarity 36.9%; Pred. No. 1.7e-55;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFTFYKLPK-RONKESGETGLELYLSDTTSRWK 71
DB 4 LLLAALLCPAGALTCYDGSQGVDFWVYVYKLPALRGSGEAAQGRQYKYLDDESSGGWD 63
QY 72 SEQLMNDTKSVLGRITLQOLYEAVASKNNTAVLIYNDGVKPK--VNSRYKYHTKGLLLW 129
DB 64 GRALINSPGAVGRSLQPLYR---SNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLL 120
QY 130 NRVOGFWLHISIPQF--PTPIPERGYDYPPTGRRNGSGICITTFKINQYEAIDSQLLVCNPN 188
DB 121 DHGCGFWLHVSNFPPPPASSAAYSWPHSACTYGTQLLCVSPFPAQFSKMGKQLTYTYPW 180
QY 189 VYSCIPATFHQELHMPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSDSFLDDIFA 248
DB 181 VVNYOLEGIFAQEFPPDLENVKGHVHVSQEPFMSNITLTSQAGAVFQSFQKFKFGDDLYS 240
QY 249 ANWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
DB 241 GWLAALGNLQVQVWHKTVGILPNSCDIWOVLNVNQIAFFPGPAGPSFNSTEDHSKNCV 300
QY 307 SOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESC 356
DB 301 SPKGP---WTCVGMNRNQGEQRGGGTILCAQLPALWKAQFQPLVKNYQPC 347

RESULT 6

US-08-639-294-2
; Sequence 2, Application US/08639294
; Patent No. 6265195
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,294
; FILING DATE: 25-Apr-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-639-294-2

Query Match 31.7%; Score 613; DB 3; Length 360;

Best Local Similarity 36.9%; Pred. No. 1.7e-55;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFTFYKLPK-RONKESGETGLELYLSDTTSRWK 71
DB 4 LLLAALLCPAGALTCYDGSQGVDFWVYVYKLPALRGSGEAAQGRQYKYLDDESSGGWD 63
QY 72 SEQLMNDTKSVLGRITLQOLYEAVASKNNTAVLIYNDGVKPK--VNSRYKYHTKGLLLW 129
DB 64 GRALINSPGAVGRSLQPLYR---SNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLL 120
QY 130 NRVOGFWLHISIPQF--PTPIPERGYDYPPTGRRNGSGICITTFKINQYEAIDSQLLVCNPN 188
DB 121 DHGCGFWLHVSNFPPPPASSAAYSWPHSACTYGTQLLCVSPFPAQFSKMGKQLTYTYPW 180
QY 189 VYSCIPATFHQELHMPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSDSFLDDIFA 248
DB 181 VVNYOLEGIFAQEFPPDLENVKGHVHVSQEPFMSNITLTSQAGAVFQSFQKFKFGDDLYS 240
QY 249 ANWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
DB 241 GWLAALGNLQVQVWHKTVGILPNSCDIWOVLNVNQIAFFPGPAGPSFNSTEDHSKNCV 300
QY 307 SOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESC 356
DB 301 SPKGP---WTCVGMNRNQGEQRGGGTILCAQLPALWKAQFQPLVKNYQPC 347

RESULT 7

US-09-861-034B-2
; Sequence 2, Application US/09861034B
; Patent No. 6569429
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861,034B

; FILING DATE: 18-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3562
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-861-034B-2

Query Match 31.7%; Score 613; DB 4; Length 360;
Best Local Similarity 36.9%; Pred. No. 1.7e-55;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFGLVGAATISCRNEKAVDFTFYKLPK-RONKESGETGLLEYLYLDSTTRGRWK 71
DB 4 LLLAALLCVFAGALTCDGSGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSEGGWRD 63
QY 72 SEQLMNDTKSYGLRTLOQLYAYASKNNNTAYLYNDGVPKP--VNTSRKYGHTKGLLLW 129
DB 64 GRALINSPEGAVGSLQPLYR--SNTSQLAFLYNDQPPQPSKAQDSSMRGHTKGVILL 120
QY 130 NRVOGFWLIHSIPF-PRIPEDGYDPTGRNQGSGICITPKYNOYEADISOLLVCNPN 188
DB 121 DHGCGFWLHVSVNFPFPASSAAYSWPHSACTYQTLTLCVSPFAQFSKMGKQLYTYVPW 180
QY 189 VYSCISIPATHOEIHPOLCTRASSBEIPRLITLQSAQGLFLHFAKSDSFLDDIFA 248
DB 181 VYVQLSGIFAEFPLENVVKGHVHVSQEPWNSSITLTSQAGAVFQSPFAKFSKFGDLYS 240
QY 249 AWAORLKHLLATWQKROKELPSNCSLPYHVYNIKAIKLSRHS--YFSSYQDHAKWCI 306
DB 241 GWLAAALGTNLQVQFHKTVGILPNSCSDIWOVLNVNQLAPPAGPSPFSENSTEDHSKWCV 300
QY 307 SOKGTNRWTCIGDLNRSFPHQAFSGGFICTONWQIYQAFQGLVLYYESC 356
DB 301 SPKGP---WTCVGMNRRNQGEQRGGTLCALPALWKAFOPLVKNYQPC 347

RESULT 8
US-09-147-915-4
; Sequence 4, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krieser, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-147-915-4

Query Match 23.7%; Score 458; DB 3; Length 275;

Best Local Similarity 36.4%; Pred. No. 1.8e-39;
Matches 95; Conservative 50; Mismatches 106; Indels 10; Gaps 5;

QY 102 AYLIYNDGVKPNY--SRKYGHTKGLLWNRVQGWFLHISIPQFP-IFEEGVYDYPPTG 158
DB 4 AFVLYNDQPPKSSKSSDSSRGHTKGVLLDQSGFWLHISVFNFPFRASSAAYSWPFGA 63
QY 159 RRGQSGICITFKYNOYEADISOLLVCNPNVYSCISIPATHOEIHPOLCTRASSSEIP 218
DB 64 QKGTGLICVSPFLTQFLDISKQLTYTYPLVYDHRLEGDFGQFPYLEEVVKGHVHVRQGP 123
QY 219 GRLLTTLQSAQGLFLHFAKSDSFLDDIFAAMAQRLKHLTETWQKQELPSNCSLP 278
DB 124 WNSSVLTLSKKGATFQSFQKFGNFGDLYSGWLAELGSLTQVQWQSSGILPNSCSGA 183
QY 279 YHYVNI--KAIKLSRHSYFSSYQDHAKWCI-SOKGTNRWTCIGDLNRSFPHQAFSGGFI 336
DB 184 QHVFDTVTQTAPPQGPAGPAFNATEDHSKWCVTPKGP---WACVGMNRRNQGEHRGGTLC 240
QY 337 TONWQIYQAFQGLVLYYESCK 357
DB 241 AQ--MLWKAFLVKAWEPC 259

RESULT 9
US-09-210-422-2
; Sequence 2, Application US/09210422A
; Patent No. 6455250
; GENERAL INFORMATION:
; APPLICANT: AGUILERA, RENATO J.
; APPLICANT: LYON, CHRISTOPHER J.
; TITLE OF INVENTION: ENDONUCLEASE COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UCLA019
; CURRENT APPLICATION NUMBER: US/09/210,422A
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/069,205
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Musca sp.
US-09-210-422-2

Query Match 11.8%; Score 229; DB 4; Length 192;
Best Local Similarity 26.5%; Pred. No. 9.1e-16;
Matches 67; Conservative 25; Mismatches 95; Indels 66; Gaps 6;

QY 6 LRTSFALLFLGFLVGAATISCRNEKAVDFTFYKLPKRONKESGETGLLEYLYLDST 65
DB 1 MATLRSLLLAALLWV-PAEALSCYDGSQPDVDFVYKLPALHSGSRDTPKGLTCKYMDQ 59
QY 66 TRSWRKSEQLMNDTKSYGLRTLOQLYAYASKNNNTAYLYNDGVPK--PVNYSRKYGH 123
DB 60 SDGWQDGVGVINSPEGAVGSLQPLYR---KNSSQLAFLYNDQPPKSSSTRDSTGHGHT 116
QY 124 KGLLWNRVQGWFLHISIPQFPPIPEGYDYPPTGRNQGSGICITFKYNOYEADISOLL 183
DB 117 KG-----KGLT 122
QY 184 VCNPNVYSCISIPATHOEIHPOLCTRASSSEI---PGRLLTTLQSAQGLFLHFAKSD 240
DB 123 YTYPLVYDHRLEGDFGQFPK---PDLETIVKQHVLPWNSSVLTLSQAGATFQSFQKFG 179
QY 241 SFLDDIFAAMWAQ 253
DB 180 KFGDLYSGWLA 192

RESULT 10
US-09-198-452A-36
; Sequence 36, Application US/09198452A


```
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 36
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-198-452A-36

Query Match
Best Local Similarity 5.0%; Score 96.5; DB 4; Length 661;
Matches 76; Conservative 59; Mismatches 157; Indels 93; Gaps 19;

QY 4 RLLRTSALLFLGLGVLGAATISCRNEEG--KAVDMFTFYKLPKQNKESGETGLE--Y 59
Db 9 RLAKRCF--LFLASFVLGSSADALTHQEAVKKNVLSHFK-----SVSGIVTIEDGV 60
QY 60 LYLDSTTRSWKSEQLMNDTKSVLGRTLQQLYEA-----YASKSNNTAYLIYNDGVKPV 114
Db 61 LNIHNNLRIOANKYVEN-----TVGOSLKLVAHGNVMVYRAKTLVCDILEY----- 108
QY 115 NYSRKYGHTKGLLWN--RVQGFVLIHSIPQPP--IPEEGYDPTGRRNGQSGICI 168
Db 109 -----VEDTDSCLTNGRFAMYPFWLGGSMITITPETIVIRKGYISTSEGPK--KDLCL 160
QY 169 TFKYNOYEAIQSLVLCNPNVYSCIPATPQHEL-----IHPQLCTRASSSEIPRL 222
Db 161 SGDYLEYSS--DLSLSIGTKITVRICRIPILFLPFSIMPELPRPINFRCGTGGFLG 219
QY 223 TTIQSQAGKFLHFAKDSFLDDIPAAWMAQRUKTHLLTETWORKRQELPNSCLPYHYV 282
Db 220 GMSYSPISRK--HFS--STFLDSFFKHGVMGFNLHC-----SQKQVENVF 263
QY 283 NIKA-----IKLSRHSYSSYQDHAKWCISQKGT-----KRWTCIGDL----- 321
Db 264 NMKSYAHLADMAEAHRYLHGDFCTHKHVNFSGEYHLSDSWETVADIPNNFM 323
QY 322 NRSHPQAFSGGFICTQNNQIQAF 346
Db 324 NTGPTRV-----DCTWNDNYPEGY 342

RESULT 11
US-08-416-870C-10
; Sequence 10, Application US/08416870C
; Patent No. 5824862
; GENERAL INFORMATION:
; APPLICANT: HIYOSHI, TORU
; APPLICANT: MINE, TOSHIKI
; APPLICANT: KASAKA, KEISUKE
; APPLICANT: TYSON, ROBERT HOW
; APPLICANT: PAGE, ANTHONY MILES JOHN
; TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
; TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
; TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALL CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P(PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-870C-10

Query Match
Best Local Similarity 4.9%; Score 95.5; DB 2; Length 426;
Matches 67; Conservative 52; Mismatches 100; Indels 149; Gaps 20;

QY 65 TTSWRKSEQLMNDTKSVLGRTLQQLYEAAYSNNNTAYLIYND-----GVPKP 113
Db 116 TSGGHDTTKIVD---SIQDRGINQVYIIGDGSQKGAIVFEIRRRGLKVAAGIPKT 172
QY 114 VN-----YSRKYV-----HTK-----GLL-LMNRVQGFVLIHS----- 140
Db 173 IDNDIPIIDRSFGDFAVBEAQAINAAHVEATSFENGIGLVKLMGRYSFIAMATLAS 232
QY 141 -----IPQFPPIPEGYDPTGRRNGQSGICITFKYNOYEAIQSLVLCNPNVYSCS 193
Db 233 RDVDCCLIPESPFLE-----GKGL-----PEFIGKRL----- 261
QY 194 IPATFHQELIHPQLCTRASSSEI---PGRLLTTLQSAQKQFLHFAKDSFLDDIPAAW 250
Db 262 -----KEIGHMVIVIAEGAGQDLLAESNEQSTTLKADSGNK-----LIQDV-GLW 305
QY 251 MAQRKTHLLTETWORKR-----QELPSN-----CSL-----PYHYVNIK 285
Db 306 ISQRIKHFAKMTNLKVIDTYMTRAVPSNADNVCCITLLAQSAVHGMAGYNGFTVG 365
QY 286 ATKLSRHSYSSYQDHAKWCISQKGTK-----NRWTCIGDLNRS-----HQAFRSG 332
Db 366 LVN-GRHTYIPFYR-----ITEKQNKVVITDRMWARLLSSTNQPSFMKDDHHEHSG 418
QY 333 GFICTQNW 340
Db 419 GEAGAMNW 426

RESULT 12
US-08-468-036-3
; Sequence 3, Application US/08468036
; Patent No. 5728806
; GENERAL INFORMATION:
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Hoekstra, Marl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that
; TITLE OF INVENTION: Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,036
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,605
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5728806and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31784
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-036-3

Query Match 4.9%; Score 95; DB 1; Length 595;
Best Local Similarity 22.5%; Pred. No. 0.55;
Matches 64; Conservative 35; Mismatches 99; Indels 86; Gaps 15;

QY 17 LFGVLGAATI-----SCRNEGKAVDWFTFYKLPKQNKESGETGLELYLDSTTR---S 68
DB 179 IFLISGVSEVNFQVQRNEIDK-IEWDFKISKITWK-----SNIKYILNSMWRPLSM 233

QY 69 W-RKSEQLMND-----TKSVLGR-----LQOYEAAYASKNNAYLIY 106
DB 234 WLRHQRIKQEDQLKSYAEQKLLGTTKEQIDPGRELNLMLHTAVQANSNNA--VS 291

QY 107 NDGVPKP-----VNSRYKGYHTKGLL----- 127
DB 292 NGQVPSQELQHLKEQSGEHNQKQDQSSFSQQQPSIFPSLSRPFANNKVIPTMPMA 351

QY 128 --LNNRVQGWLIHSIPQPIPEEGYDYPPTGRNGOSGICITPKYNOYEAIDSQLVC 185
DB 352 NVFMSNPOLFATMGQP-FAPFP---FMLPLTNNNSANPIPTVPPN-FNAPNPMAFG 406

QY 186 NPNVYSCIPATFHQELIHPQLCTRAS--SSEIPGRLLTTIQS 227
DB 407 VPMNHLGPAV-SQPFSLPAPLPRDSGYSSSSPGQLDLILNS 449

RESULT 13
US-08-376-843-3
; Sequence 3, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Hoeckstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; TITLE OF INVENTION: that Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/376,843
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,605
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5846764and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31784
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-376-843-3

Query Match 4.9%; Score 95; DB 2; Length 595;
Best Local Similarity 22.5%; Pred. No. 0.55;
Matches 64; Conservative 35; Mismatches 99; Indels 86; Gaps 15;

QY 17 LFGVLGAATI-----SCRNEGKAVDWFTFYKLPKQNKESGETGLELYLDSTTR---S 68
DB 179 IFLISGVSEVNFQVQRNEIDK-IEWDFKISKITWK-----SNIKYILNSMWRPLSM 233

QY 69 W-RKSEQLMND-----TKSVLGR-----LQOYEAAYASKNNAYLIY 106
DB 234 WLRHQRIKQEDQLKSYAEQKLLGTTKEQIDPGRELNLMLHTAVQANSNNA--VS 291

QY 107 NDGVPKP-----VNSRYKGYHTKGLL----- 127
DB 292 NGQVPSQELQHLKEQSGEHNQKQDQSSFSQQQPSIFPSLSRPFANNKVIPTMPMA 351

QY 128 --LNNRVQGWLIHSIPQPIPEEGYDYPPTGRNGOSGICITPKYNOYEAIDSQLVC 185
DB 352 NVFMSNPOLFATMGQP-FAPFP---FMLPLTNNNSANPIPTVPPN-FNAPNPMAFG 406

QY 186 NPNVYSCIPATFHQELIHPQLCTRAS--SSEIPGRLLTTIQS 227
DB 407 VPMNHLGPAV-SQPFSLPAPLPRDSGYSSSSPGQLDLILNS 449

RESULT 14
US-09-409-180A-1
; Sequence 1, Application US/09409180A
; Patent No. 6444802
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: White, David
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 22196, A No. 6444802el Human Amino peptidase
; FILE REFERENCE: 5800-59
; CURRENT APPLICATION NUMBER: US/09/409,180A
; CURRENT FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-409-180A-1

Query Match 4.9%; Score 95; DB 4; Length 704;
Best Local Similarity 18.7%; Pred. No. 0.71;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISCRNE-----EGKAVDWFTFYKLPKQNKES---SGETGLELYLDSTTRSWRKS 72
```

[illegible]

RESULT 15

```

US-09-071-035-276
; Sequence 276, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 276:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-276

```

Query Match	4.7%	Score 90.5;	DB 4;	Length 384;
Best Local Similarity	21.8%;	Pred. NO. 0.82;		
Matches	82;	Conservative	52;	Mismatches 111;
				Indels 111; Gaps 21;

```

Db      42  EGKTV-----OEAQRQLAETGATNRKPIAEDKT-----QTINGEDGVVSFSL----- 83
Qy      92  EAVASKNNTAYLLYNDGVPKPVNYSRKYGHTKGLLWNRVQGFLLHHSIPOFPPIPESG 151
Db      84  ASKDQORDKAYLFVEAEAEVW--KEKASNLVILPVODPQGQSLTH--IHLYPKNEENA 140
Qy      152  YDVPPTGR--NGOSGICITFKYNQYRAIDSQLLAVCNPNVYSCSIPATF--HOELIHMPO 207
Db      141  YDLPPLEKTVLDKQOQ-----FNQGEHINYQL-----TTQIPANILGQOEF---R 182
Qy      208  LCTWASS-----EIPGRLLTTLQSOQOKF---LHFAKSD--SFLDDIFAAWMA 252
Db      183  LSDKADPTLLPESIEBKVAGKTVTVTYTLTQKHGFTLDFSIKDLQNFANQMTVSVQ 242
Qy      253  QRLK-----THLLTE-----TWQRKQBLPS---NCSLPYHVYNTKA 286
Db      243  MRLEKTAEPDTAINNEGQVTDGHTLTKRATVRTGGKSFVKVDSENAKITLPEAVF---I 299
Qy      287  ICLSRHSYFSYQDHAKW-----CISQKGTKN-----RW-----TCIGDLNRS 324
Db      300  VKNQAGEYLNETANGYRQWEKALAKFTSNQAGEFVKGKRWPVLLGRNLTKRUSSE 359
Qy      325  PHQAFRSGG--FICTQ 338
Db      360  SNRNSFYGGKKFLCNE 375

```

Search completed: October 14, 2004, 00:33:37
Job time : 33.6329 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:31:21 ; Search time 67.7848 Seconds
(without alignments)
1694.809 Million cell updates/sec

Title: US-10-790-589-4
Perfect score: 1933
Sequence: 1 MMARLRTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pcp:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pcp:*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1933	100.0	357	9	US-09-949-434-4
2	1933	100.0	357	12	US-10-240-709-4
3	1933	100.0	357	16	US-10-790-589-4
4	1926	99.6	361	16	US-10-670-863-3
5	1293.5	66.9	354	9	US-09-949-434-2
6	1293.5	66.9	354	12	US-10-240-709-2
7	1293.5	66.9	354	16	US-10-790-589-2
8	1293.5	66.9	354	16	US-10-670-863-1
9	613	31.7	360	15	US-10-408-167A-2
10	580	30.0	348	16	US-10-408-765A-1164
11	98.5	5.1	391	12	US-10-282-122A-69762
12	97.5	5.0	697	16	US-10-467-534-39
13	96.5	5.0	661	15	US-10-289-762-36
14	95	4.9	681	15	US-10-362-226-1
15	95	4.9	687	10	US-09-842-758-24

16	95	4.9	687	12	US-10-174-333-24	Sequence 24, Appl
17	95	4.9	704	10	US-09-842-758-26	Sequence 26, Appl
18	95	4.9	704	12	US-10-174-333-26	Sequence 26, Appl
19	95	4.9	704	12	US-09-972-211-98	Sequence 98, Appl
20	95	4.9	704	12	US-09-833-782-2	Sequence 2, Appl
21	95	4.9	704	12	US-10-096-625-98	Sequence 98, Appl
22	95	4.9	704	14	US-10-192-207-1	Sequence 1, Appl
23	95	4.9	704	14	US-10-192-445-1	Sequence 1, Appl
24	95	4.9	714	16	US-10-311-035-4	Sequence 4, Appl
25	95	4.9	884	12	US-10-424-599-214353	Sequence 214353,
26	94	4.9	537	16	US-10-322-281-368	Sequence 368, App
27	93	4.8	526	12	US-10-382-122A-76916	Sequence 76916, A
28	92.5	4.8	272	16	US-10-437-963-152834	Sequence 22, Appl
29	92.5	4.8	638	10	US-09-842-758-22	Sequence 22, Appl
30	92.5	4.8	638	12	US-10-174-333-22	Sequence 22, Appl
31	92	4.8	645	16	US-10-322-281-364	Sequence 364, App
32	91	4.7	653	16	US-10-437-963-173197	Sequence 173197,
33	91	4.7	1063	15	US-10-161-493-64	Sequence 64, Appl
34	91	4.7	1063	15	US-10-161-493-66	Sequence 66, Appl
35	90.5	4.7	384	12	US-10-206-576-276	Sequence 276, App
36	90.5	4.7	430	12	US-10-206-576-274	Sequence 274, App
37	90.5	4.7	695	15	US-10-312-273-91	Sequence 91, Appl
38	90	4.7	418	12	US-10-424-599-189389	Sequence 189389,
39	89.5	4.6	476	16	US-10-661-809-11	Sequence 11, Appl
40	89	4.6	527	9	US-09-738-626-4803	Sequence 4803, Ap
41	89	4.6	607	15	US-10-108-260A-3710	Sequence 3710, Ap
42	88.5	4.6	624	12	US-09-972-211-24	Sequence 24, Appl
43	88.5	4.6	624	12	US-10-096-625-24	Sequence 24, Appl
44	88.5	4.6	704	12	US-09-972-211-12	Sequence 12, Appl
45	88.5	4.6	704	12	US-09-972-211-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-949-434-4
; Sequence 4, Application US/09949434
; Patent No. US20020028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-434-4

Query Match	100.0%	Score	1933	DB	9	Length	357
Best Local Similarity	100.0%	Pred. No.	1.3e-186				
Matches	357	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MMARLRTSFALLFLGLFGVLCGAVTISCRNEGKAVDWFTFYKLPKRNKESGETGLEYL	60				
Db	1	MMARLRTSFALLFLGLFGVLCGAVTISCRNEGKAVDWFTFYKLPKRNKESGETGLEYL	60				
QY	61	YLDSTTRSRKSEQLMNDTKSVLGRITLQQLYAYASKNNNTAYLYNDGVKPVNYSRY	120				
Db	61	YLDSTTRSRKSEQLMNDTKSVLGRITLQQLYAYASKNNNTAYLYNDGVKPVNYSRY	120				
QY	121	GHTKGLLLNRRVQGFLLHISIPQFPPIPEEGYDYPPTGRNGQSGICITFKKNQVEAIDS	180				
Db	121	GHTKGLLLNRRVQGFLLHISIPQFPPIPEEGYDYPPTGRNGQSGICITFKKNQVEAIDS	180				
QY	181	QLLVNPNVYSCISIPATHQELIHPQLCTRASSSEIPGRLLITLQSAQGGKFLHFAKSD	240				

Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTLTQSAQKQFLHFAKSD 240
QY 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300
Db 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 2
US-10-790-589-4
; Sequence 4, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-709-4

Query Match 100.0%; Score 1933; DB 12; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.3e-186; Indels 0; Gaps 0;
Matches 357; Conservative 0; Mismatches 0

QY 1 MVARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKQNKESGETGLEYL 60
Db 1 MVARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKQNKESGETGLEYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120

QY 121 GHTKGLLLNRRVQGFLLHSIPQPPPEEGYDYPPTGRNGQSGICITFKYNOYEADSD 180
Db 121 GHTKGLLLNRRVQGFLLHSIPQPPPEEGYDYPPTGRNGQSGICITFKYNOYEADSD 180

QY 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTLTQSAQKQFLHFAKSD 240
Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTLTQSAQKQFLHFAKSD 240

QY 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300
Db 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 3
US-10-790-589-4
; Sequence 4, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match 100.0%; Score 1933; DB 16; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.3e-186; Indels 0; Gaps 0;
Matches 357; Conservative 0; Mismatches 0

QY 1 MVARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKQNKESGETGLEYL 60
Db 1 MVARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKQNKESGETGLEYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120

QY 121 GHTKGLLLNRRVQGFLLHSIPQPPPEEGYDYPPTGRNGQSGICITFKYNOYEADSD 180
Db 121 GHTKGLLLNRRVQGFLLHSIPQPPPEEGYDYPPTGRNGQSGICITFKYNOYEADSD 180

QY 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTLTQSAQKQFLHFAKSD 240
Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTLTQSAQKQFLHFAKSD 240

QY 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300
Db 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 4
US-10-670-863-3
; Sequence 3, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohawa, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 3
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-670-863-3

Query Match 99.6%; Score 1926; DB 16; Length 361;
Best Local Similarity 99.7%; Pred. No. 6.5e-186; Indels 0; Gaps 0;
Matches 356; Conservative 0; Mismatches 1

QY 1 MVARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKQNKESGETGLEYL 60

Db 5 MMARLLRTSFALLFLGLFGLVGAANTISCRNEBEGKAVDWFTFYKLPKRONKESGETGLEYL 64
 QY 61 YLDSTTRSRWKSQQLMNDTKSVLGRITLOQLYEAYASKNNTAYLIYNDGVKPVNYSKY 120
 Db 65 YLDSTTRSRWKSQQLMNDTKSVLGRITLOQLYEAYASKNNTAYLIYNDGVKPVNYSKY 124
 QY 121 GHTKGLLLNWRVQGFLLIHSIPQPPPIPEGGYDPTGRRNGSGICITFKYNOYEADS 180
 Db 125 GHTKGLLLNWRVQGFLLIHSIPQPPPIPEGGYDPTGRRNGSGICITFKYNOYEADS 184
 QY 181 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLITLLOSAQOGKFLHFAKSD 240
 Db 185 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLITLLOSAQOGKFLHFAKSD 244
 QY 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPWHVYNIKAIKLSRHSYFSSYQD 300
 Db 245 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPWHVYNIKAIKLSRHSYFSSYQD 304
 QY 301 HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
 Db 305 HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 361

RESULT 5

US-09-949-434-2

; Sequence 2, Application US/09949434
 ; Patent No. US2002028495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan R.
 ; APPLICANT: Krieser, Ronald J.
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
 ; FILE REFERENCE: DC-0137
 ; CURRENT APPLICATION NUMBER: US/09/949,434
 ; PRIOR FILING DATE: 2001-09-07
 ; PRIOR APPLICATION NUMBER: 09/574,942
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 US-09-949-434-2

Query Match 66.9%; Score 1293.5; DB 9; Length 354;
 Best Local Similarity 66.4%; Pred. No. 6.7e-122;
 Matches 23; Conservative 45; Mismatches 72; Indels 3; Gaps 2;
 QY 1 MMARLLRTSFALLFLGLFGLVGAANTISCRNEBEGKAVDWFTFYKLPKRONKESGETGLEYL 60
 Db 1 MTAKPLRTVLSLFFALLSGVLGTPETSCRNEYGEAVDWFTFYKLPKRTSKASEAGLQYL 60
 QY 61 YLDSTTRSRWKSQQLMNDTKSVLGRITLOQLYEAYASKNNTAYLIYNDGVKPVNYSKY 120
 Db 61 YLDSTRTQTNKSLYLINSTRSALGRITLQHLDYTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
 QY 121 GHTKGLLLNWRVQGFLLIHSIPQPPPIPEGGYDPTGRRNGSGICITFKYNOYEADS 180
 Db 120 GHAKGLLVNRTQGFLLIHSVPKPPV--HGYYPTSGRRYGTGICITFGYSQFBEIDF 177
 QY 181 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLITLLOSAQOGKFLHFAKSD 240
 Db 178 QLLVCPNVYSCSIPSTFHWKLLYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
 QY 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPWHVYNIKAIKLSRHSYFSSYQD 300
 Db 238 FYDDIFTGWAQKLAQKTHLLAQTWQKKQELPNSCLPWHVYNIKSIKSVTSKYSFSSRQD 297
 QY 301 HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
 Db 298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRYGFCCK 354

RESULT 6

US-10-240-709-2
 ; Sequence 2, Application US/10240709
 ; Publication No. US20030212023A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan R.
 ; APPLICANT: Krieser, Ronald J.
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
 ; FILE REFERENCE: DC-0154
 ; CURRENT APPLICATION NUMBER: US/10/240,709
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: 09/541,840
 ; PRIOR FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: 09/574,942
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 US-10-240-709-2

Query Match 66.9%; Score 1293.5; DB 12; Length 354;
 Best Local Similarity 66.4%; Pred. No. 6.7e-122;
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;
 QY 1 MMARLLRTSFALLFLGLFGLVGAANTISCRNEBEGKAVDWFTFYKLPKRONKESGETGLEYL 60
 Db 1 MTAKPLRTVLSLFFALLSGVLGTPETSCRNEYGEAVDWFTFYKLPKRTSKASEAGLQYL 60
 QY 61 YLDSTTRSRWKSQQLMNDTKSVLGRITLOQLYEAYASKNNTAYLIYNDGVKPVNYSKY 120
 Db 61 YLDSTRTQTNKSLYLINSTRSALGRITLQHLDYTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
 QY 121 GHTKGLLLNWRVQGFLLIHSIPQPPPIPEGGYDPTGRRNGSGICITFKYNOYEADS 180
 Db 120 GHAKGLLVNRTQGFLLIHSVPKPPV--HGYYPTSGRRYGTGICITFGYSQFBEIDF 177
 QY 181 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLITLLOSAQOGKFLHFAKSD 240
 Db 178 QLLVCPNVYSCSIPSTFHWKLLYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
 QY 241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPWHVYNIKAIKLSRHSYFSSYQD 300
 Db 238 FYDDIFTGWAQKLAQKTHLLAQTWQKKQELPNSCLPWHVYNIKSIKSVTSKYSFSSRQD 297
 QY 301 HAKWCISQKGTNRWTCIGDLNRSHPQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
 Db 298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRYGFCCK 354

RESULT 7

US-10-790-589-2
 ; Sequence 2, Application US/10790589
 ; Publication No. US20040142376A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan R.
 ; APPLICANT: Krieser, Ronald J.
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
 ; FILE REFERENCE: DC-0154
 ; CURRENT APPLICATION NUMBER: US/10/790,589
 ; CURRENT FILING DATE: 2004-03-01
 ; PRIOR APPLICATION NUMBER: 09/541,840
 ; PRIOR FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: 09/574,942
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 354


```

; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-589-2

Query Match      66.9%; Score 1293.5; DB 16; Length 354;
Best Local Similarity 66.4%; Pred. No. 6.7e-122;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MMARLRTSFALLFLGLFGLVGAATISCRNEEGKAVDWFYFKLPKRNKESGEGTGLYL 60
Db 1 MTAKPLRTVLISLLFFALLSGVLGTPEISCRNEYGEAVDWFYFKLPKRTSKASEAGLYL 60

Qy 61 YLDTSTRSWKSQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNSRY 120
Db 61 YLDTSTRQWNKSLYLINSTRSALGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNSRY 119

Qy 121 GHTKGLLNNRVQGVFWLHISIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180
Db 120 GHAKGLLVNRTQGVFWLHISVPKPPV--HGVEYPTSGRRYQGTGICITFGVSQPEEIDF 177

Qy 181 QLLVCNPNVYSCSDPATTHQELIHPQLCTRASSEIFGRLLITLQSAQKQFLHFAKSD 240
Db 178 QLLVQPNYISCFIPSTFTHMKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

Qy 241 SFLDDIFAANWAORLKTLLTETWQRKQELPNSCSLPHVYVNIKAIKLSRHSYFSSYQD 300
Db 238 FYTDDIFTGWIATQAKLKTLLTETWQRKQELPNSCSLPHVYVNIKAISIGVTSKYSFSSRQD 297

Qy 301 HAKWCISQKGTNRWTCIGDLNRSRPHQAFRSGGFTCTQNWQIYQAFQGLVLYYESCK 357
Db 298 HSKWCVSIGSANSRWTCTIGDLNRSRPHQALRGGGFTCTKNHYIYQAFHKLRYLRYGCK 354

```

```

RESULT 8
US-10-670-863-1
; Sequence 1, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-670-863-1

```

```

Query Match      66.9%; Score 1293.5; DB 16; Length 354;
Best Local Similarity 66.4%; Pred. No. 6.7e-122;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MMARLRTSFALLFLGLFGLVGAATISCRNEEGKAVDWFYFKLPKRNKESGEGTGLYL 60
Db 1 MTAKPLRTVLISLLFFALLSGVLGTPEISCRNEYGEAVDWFYFKLPKRTSKASEAGLYL 60

Qy 61 YLDTSTRSWKSQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNSRY 120
Db 61 YLDTSTRQWNKSLYLINSTRSALGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNSRY 119

Qy 121 GHTKGLLNNRVQGVFWLHISIPQPPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDS 180
Db 120 GHAKGLLVNRTQGVFWLHISVPKPPV--HGVEYPTSGRRYQGTGICITFGVSQPEEIDF 177

Qy 181 QLLVCNPNVYSCSDPATTHQELIHPQLCTRASSEIFGRLLITLQSAQKQFLHFAKSD 240

```

```

Db 178 QLLVQPNYISCFIPSTFTHMKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Qy 241 SFLDDIFAANWAORLKTLLTETWQRKQELPNSCSLPHVYVNIKAIKLSRHSYFSSYQD 300
Db 238 FYTDDIFTGWIATQAKLKTLLTETWQRKQELPNSCSLPHVYVNIKAISIGVTSKYSFSSRQD 297
Qy 301 HAKWCISQKGTNRWTCIGDLNRSRPHQAFRSGGFTCTQNWQIYQAFQGLVLYYESCK 357
Db 298 HSKWCVSIGSANSRWTCTIGDLNRSRPHQALRGGGFTCTKNHYIYQAFHKLRYLRYGCK 354

```

RESULT 9

```

US-10-408-167A-2
; Sequence 2, Application US/10408167A
; Publication No. US20030219428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/408,167A
; FILING DATE: 04-Apr-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/861034
; FILING DATE: 18-May-2001
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, David W.
; REGISTRATION NUMBER: NONE
; REFERENCE/DOCKET NUMBER: P1024DIC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1739
; TELEFAX: 650/952-3881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-408-167A-2

```

```

Query Match      31.7%; Score 613; DB 15; Length 360;
Best Local Similarity 36.9%; Pred. No. 5.2e-53;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

Qy 13 LFLGLGVIGVGAATISCRNEEGKAVDWFYFKLPK-RNKESGEGTGLVLYLDSTTSWRK 71
Db 4 LLLAALLCVFAGALTCYGDGQFVDFVYVYKLPALRGSGEAAQGRQYQYLDSSGWRD 63

Qy 72 SEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPV--VNSRYKYGHTKGLLW 129
Db 64 GRALINSPGAVGRSLQPIYR---SNTSLAFLIYNDQPPQPSKAQDSSMRGHTKGVLL 120

Qy 130 NRYQGVFWLHISIPQF-PPPIPEEGYDYPPTGRRNGQSGICITFKINQYEAIDSLVCNPN 188
Db 121 DHDDGGFWLHVSFVNFPPPPASSAAYSWPHSACTYGTLLCVSPFPFAQFSKMGKOLTYTPW 180

Qy 189 VYSCSDPATTHQELIHPQLCTRASSEIFGRLLITLQSAQKQFLHFAKSDSFLDDIFA 248

```

Db 181 VNYQLEGIFAQEPDLENVKGHHVSGEPWNSITLTSQAGAVFQSKFAKTSKFGDDLYS 240
QY 249 AWMAQLKTHLLTETWQRKQELPNSCSLPYHYNYKAIKLSRHS--YFSSYQDHAKWCI 306
Db 241 GWLAAALGTNLQVFWHTVGLPNSCDIWOVLNVNQIAPFGPAGPSFNSTEDHSKWCY 300
QY 307 SQKGTNRWTCIGDLNRSFQAFRSGGFICTQNWQIYQAFQGLVLYYESC 356
Db 301 SPKGP---WTCVGMNRNQEGRGGTLCAQLPALWKAQFPLVKNYQPC 347

RESULT 10

US-10-408-765A-1164

; Sequence 1164, Application US/10408765A

; Publication No. US20040101874A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Faby, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1164

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1164

Query Match

Best Local Similarity 30.0%; Score 580; DB 16; Length 348;

Matches 127; Conservative 56; Mismatches 143; Indels 24; Gaps 7;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFYFKLPK-RQNKESGEGTGLYLYLDSTTRSWRK 71
Db 4 LLLAALLCVPAAGALTCVSGSQPVDWFVVKLPALRGSGEAAQKGLQKYLDESSGGRD 63
QY 72 SEQLMNDTKSVLGRTLQQLYEAVASKNNYALYNDGVKPP--VNYSRKYGHTKGLLLW 129
Db 64 GRALINSPEGAVERSGQPLTR---SNTSLAFLYNDOPPPSKAQQSSMRGHTKGVL 120
QY 130 NRVQGFVLHISIPQF-PPIPEEGYDYPPTGRRNGQSGICITFKYNOYEAIDSOLLVCNPN 188
Db 121 DHDDGFVLVHVSFVNFPPASSAAYSWPHSACTYGTLLC-----KQLYTYPW 168
QY 189 VYSCSPATPHOLHMPQLCTRASSEIPGRLLTLOSAGQKFLHFAKSDSFLDDIFA 248
Db 169 VNYQLEGIFAQEPDLENVKGHHVSGEPWNSITLTSQAGAVFQSKFAKTSKFGDDLYS 228
QY 249 AWMAQLKTHLLTETWQRKQELPNSCSLPYHYNYKAIKLSRHS--YFSSYQDHAKWCI 306
Db 229 GWLAAALGTNLQVFWHTVGLPNSCDIWOVLNVNQIAPFGPAGPSFNSTEDHSKWCY 288
QY 307 SQKGTNRWTCIGDLNRSFQAFRSGGFICTQNWQIYQAFQGLVLYYESC 356
Db 289 SPKGP---WTCVGMNRNQEGRGGTLCAQLPALWKAQFPLVKNYQPC 335

RESULT 11

US-10-282-122A-69762

; Sequence 69762, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69762
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69762

Query Match

Best Local Similarity 5.1%; Score 98.5; DB 12; Length 391;

Matches 63; Conservative 33; Mismatches 110; Indels 57; Gaps 11;

QY 15 LGLFVLGAATISCRNEEGKAVDWFYFKLPKQNKESGET-----GLEVLYL-DS 64
Db 10 LGL-GANGAATVQLAKAGVDVIGVDYAPPHQTQSSHGDTITRLSVGSGPQLPLVRS 68
QY 65 TTRSWRKSEQLMNDTK---SVLGRTLQQLYEAYASKNNYALYNDGVKPPVNYSRKY 120
Db 69 SHRIWELEALSCELSFEQCGVLVMTSSPSYDAADA-----DDFTHKTIALARAY 118
QY 121 GHTKGLLLNVRVQFVLIHSIPQFPI-----PEGYDYPPTGRRNGQSGICITFK 171
Db 119 GVEHQVIMATAIR-----ERFPQFAPVLDTAIGYFEPGFGFVRPER-----CIGVQ 164
QY 172 YNQYEAIDSOLL---VCNPNVYSCSIPATFHQELHMPQLCTRAS--SSEIPGRLLTTL 225
Db 165 LRLAQAQGLARLTHETVTHLQYGEQVRIITRKSGSIADKVVVSVAGMWSTELLGAPFSDL 224
QY 226 QSAQKQKFLHFAKSDSFLDDIFA 248
Db 225 LRVCRQKLFWRERQ---DAVEA 244

RESULT 12

US-10-467-534-39

; Sequence 39, Application US/10467534

; Publication No. US20040131625A1

; GENERAL INFORMATION:

```

; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Veilant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-39

Query Match      5.0%; Score 97.5; DB 16; Length 697;
Best Local Similarity 24.9%; Pred. No. 2.1;
Matches 62; Conservative 25; Mismatches 109; Indels 53; Gaps 12;

QY 114 VVSEK-----YHTKGLLWN-RVQGF-WLIH---SIQFPPIPEGVDYPTG 158
DB 85 VNYRKILICDVEYEDTSCLLNGRCSLYPWFIGSTIYSPSSIIHKGYISTSEG 144
QY 159 RRMGSGICITPKYQYBAIDSQLVCNPNVYSCSIPATFHQEL-----IHPOLCTRA 212
DB 145 ---PQHICLSGDLKYSS-DSVLSNGPSRLSICNTFVLLLPQISIMPEIPKPIITFRG 200
QY 213 SSSEIPRELLTTLQAAQOKFLHFAKSDSFLDDIFAAMWAQELKTHLLTETWQRKREL 272
DB 201 GSGFLGSLVGSYSPISK--HCSITLFLDGFPHKIG-----LGNMRFSSQNP 250
QY 273 SNCSLPYHYNIKAIKLSRHSYFSSYODHAKWCISOKGTNRTWTCIGDLNRPQAFRSG 332
DB 251 SN-----AINIK-----SYAH-RLAIDSSGAKRYRLHGDFTFSKERAHLAG 292
QY 333 GFICTQNWQ 341
DB 293 EPHLSDSWE 301

RESULT 13
US-10-289-762-36
; Sequence 36, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 36
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-36

Query Match      5.0%; Score 96.5; DB 15; Length 661;
Best Local Similarity 19.7%; Pred. No. 2.5;
Matches 76; Conservative 59; Mismatches 157; Indels 93; Gaps 19;

QY 4 RLRTSFALLFLGLVGLGATISCRNEEG--KAVDMFTFKLPKQNKSGTGLE--Y 59
DB 9 RLMKRCF--LFLASFLVLMGSSADALHQAQVKKKNSTLSHFK-----SVSGIVTIEDGV 60
QY 60 LYLDSTTRSRKSBQLMNDTKSVLGRTLQQLYEA-----YAKSNNTAYLIYNDGVPKPV 114

; APPLICANT: Daichi Pharmaceutical Co., Ltd
; TITLE OF INVENTION: No. US20040014939A1
; Publication No. US20040014939A1
; GENERAL INFORMATION:
; APPLICANT: DAICHI PHARMACEUTICAL CO., LTD
; TITLE OF INVENTION: No. US20040014939A1
; FILE REFERENCE: GP01-1023
; CURRENT APPLICATION NUMBER: US/10/362,226
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: JP P2000-252503
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-362-226-1

Query Match      4.9%; Score 95; DB 15; Length 681;
Best Local Similarity 18.7%; Pred. No. 3.7;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISCRNE---EGKAVDMFTFKLPKQNKKE---SGETGLVLYLDSTTRSRKS 72
DB 314 ILNLKKCKCKRGFEYDGGKINAWDLYYNTQTEELKYSIDQBFLEKFFIEVVT----- 367
QY 73 EOLMNDTKSVLGRTLQQLYEAASKNNTAYLI-----YNDGVKPK--PVNYSR 118
DB 368 EGLNTYQELGLSFEQMDAHWNKSVTLVYVKDKATGEVLGQFLDLYPRGKYNHAA 427
QY 119 KYCHTKGLLLMNRVQGFVLIHSIPQPPPIPEGVDYPTGRRNGQSGICITFKYNOYEA 178
DB 428 CFGLQPGCLL-----PDGSRMVAALVNF----- 453
QY 179 DSQLLVCNPNVYSCSIPATFHQELIH-MPOLCTRASS-----ETPGRLLTT- 224
DB 454 -SQVAGRESLLRHDEVRTYFHFHGVHMHQICQTDFAFGSGTNVETDFVEFVSQMLENW 512
QY 225 LQSAQOQKFL--HFAKSDSFLDDIFAAMWAQRI--KTHLLT---ETWQRKRELPSNCSLP 278
DB 513 VMDVDSLRSLRKHKYGKSGPIADDDLEKLVASRLVNTGLTLRLQIVLSKVDQSLHTWTS- 571
QY 279 YHYNIKAIKLSRHSYFSSYQDHAKWC 305
DB 572 -----DAASEYAKYC 581

```

```

RESULT 15
US-09-842-758-24
; Sequence 24, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkars, Richard A
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Zernhusen, Bryan D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E
; APPLICANT: Gangolli, Baha A
; APPLICANT: Smithson, Glennda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Grosse, William M
; APPLICANT: Edward, Szekeres S
; APPLICANT: Alsobrook II, John P
; TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-758-24

```

```

Query Match 4.9%; Score 95; DB 10; Length 687;
Best Local Similarity 18.7%; Pred. No. 3.7;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISGNE-----EGKAVDMFTYKLPKQNK-----SGETGLYLXLDSTRSRKS 72
DB 320 ILMLKKECKDRGFYDGGKINAWDLYYYMTQTTELKYSIDQEPFKYFPIEVVT----- 373

```

```

QY 73 BQLMNDTKSVLGRITLQOLYAYASKSNNTAYLI-----YNDGVPK--PVNYSR 118
DB 374 EGLNTYQELLGLSFEQMTDAHVNKSVTLTYTKDKATGEVLGQFYLDLPRKGYNHAA 433
QY 119 KYGHTKGLLWNRVQGFWLIHSIPQFPPIPEEGVDYPPPTGRRNGQSGICITFFKYNQYEAI 178
DB 434 CFGIQQGCLL-----PDGSEWVAVALVNF----- 459
QY 179 DSQLLVCNPNVYSCSIPATFHQLIH-MPOLCTRASS-----EIPGRLLTT- 224
DB 460 -SQPVAGRPSSLRRHDEVRTYFHEFGVHMHCQAQTFARFSGTNVETDFEVPQMLENW 518
QY 225 LQSAQGQKFL--HFAKSDSFLDDIFAAWMAORL-KTHLIT---ETWORKKEQLPNSCISLP 278
DB 519 VMDVDSLRLRSKHYKDGSPITADDLLEKLVASRLVNTGLLTROIIVLSKVDQSLHTNTSL- 577
QY 279 YHVYNIKAIKLSRHSYFSSYQDHAKWC 305
DB 578 -----DAASEYAKYC 587

```

Search completed: October 14, 2004, 00:45:59
Job time : 69.7848 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:08:06 ; Search time 29,1224 Seconds
(without alignments)
1179.175 Million cell updates/sec

Title: US-10-790-589-4
Perfect score: 1933
Sequence: 1 MWARLLRTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1297.5	67.1	356	JC7131	deoxyribonuclease
2	613	31.7	360	JE0206	deoxyribonuclease
3	593	30.7	364	JE0205	deoxyribonuclease
4	580	30.0	348	T45071	hypothetical prote
5	490.5	25.4	375	T19038	hypothetical prote
6	413	21.4	516	S40996	hypothetical prote
7	319	16.5	238	S44793	F09G8.2 protein -
8	117	6.1	344	A44164	secreted glycoprot
9	97.5	5.0	697	H71525	probable outer mem
10	96.5	5.0	704	C81542	conserved hypothet
11	95	4.9	970	S63059	hypothetical prote
12	94.5	4.9	2295	B71621	probable membrane
13	94	4.9	523	DEFF06	glucose-6-phosphat
14	93.5	4.8	394	E81286	probable polysacch
15	93	4.8	526	F82873	hypothetical prote
16	93	4.8	6689	S55024	nebulin, skeletal
17	92	4.8	524	A47740	glucose-6-phosphat
18	92	4.8	641	F90832	terminase large su
19	92	4.8	641	C90901	terminase large su
20	92	4.8	641	H85689	terminase large su
21	92	4.8	709	C64057	Iga-specific metal
22	91.5	4.7	305	S73645	probable lipoprote
23	90.5	4.7	695	B72129	probable outer mem
24	90.5	4.7	695	F86493	probable outer mem
25	89.5	4.6	533	D86756	prophage p12 prote
26	89	4.6	831	T05771	beta-galactosidase
27	88.5	4.6	585	T22649	hypothetical prote
28	88.5	4.6	1456	A36563	mannose receptor p
29	88	4.6	424	1 MWV2K3	K3 protein - vacci

30	88	4.6	641	1 JVBPAL	DNA-packaging prot
31	87.5	4.5	700	2 B81682	conserved hypothet
32	87.5	4.5	877	2 T03098	p97 protein - Toxo
33	87	4.5	685	1 A47102	system b(0,+) amin
34	87	4.5	777	2 T25761	hypothetical prote
35	86	4.4	524	2 S76140	hypothetical prote
36	86	4.4	704	2 A45985	microsomal endopep
37	85.5	4.4	508	2 S74537	anthranilate synth
38	85.5	4.4	521	2 T11166	CDPdiacylglycerol-
39	85.5	4.4	587	2 T07634	pollen-specific pr
40	85.5	4.4	951	1 SYECVT	valine-tRNA ligase
41	85.5	4.4	951	2 C91283	valine tRNA synth
42	85.5	4.4	951	2 E86124	valine tRNA synth
43	85.5	4.4	951	2 AC1061	valine-tRNA ligase
44	85.5	4.4	954	2 G64121	valine-tRNA ligase
45	85.5	4.4	1307	2 AH0446	probable exported

ALIGNMENTS

RESULT 1

JC7131

deoxyribonuclease I (EC 3.1.21.1) - rat (strain Wistar)

N:Alternate names: deoxyribonuclease II-like acid deoxyribonuclease

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Dec-2002

C/Accession: JC7131

R:Tanuma, S.; Shiohawa, D.

Biochem. Biophys. Res. Commun. 265, 395-399, 1999

A:Title: Cloning of a cDNA encoding a rat DNase II-like acid DNase.

A:Reference number: JC7131; MUID:20025354; PMID:10558878

A:Accession: JC7131

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-356 <TAN>

A:Cross-references: GB:AF178974; NID:g6470130; PID:g6470131

C:Superfamily: deoxyribonuclease II

C:Keywords: hydrolase

Query Match

Best Local Similarity 67.1%; Score 1297.5; DB 2; Length 356;

Matches 236; Conservative 46; Mismatches 73; Indels 1; Gaps 1;

Qy	1	MWARRLLRTSFALLFLGLFGVIGATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL	60
Db	1	MTAQPLKAALPLFLFVALSGVLGTPVISCINEDGKAVDWFYKLPRTSRGGTGMGLDYL	60
Qy	61	YLDSTTRSWKSKQLMNDTKSVLGRITLQLYEAYASKNNNTAVILYNDGVPKPNYSRKY	120
Db	61	YLDSTTRSWKSKHLLNSRSSISGRTLEQYEAHNAK-NDTAVILYNDVAPASVNSGNY	119
Qy	121	GHTKGLLLNNRVQGFMLIHSPQPIPERGYDYPPTGRRNGSGICITFKYNOYEALDS	180
Db	120	GHAAGLLVNNRVQGFMLIHSPKFPVPEKGYEYPSGSGRYAQSGLCITLIKSYETIDS	179
Qy	181	QLLVCHPNVSCSIPATFTHQELHMPOLCTRASSETIPGRLLTTLQSAQGKFLHFAKSD	240
Db	180	QLLVFQPNVSCSIPATFTHQELHMPOLCTRASSETIPGRLLTTLQSAQGLNLFHFAKST	239
Qy	241	SELDIDFAAWAQRKLTLLTETWQRKQELPNSCSLPYHYVNIKALKSRHSYFSYQD	300
Db	240	FYTDDIDFAAWIAOKLVKLLVESWQRKNHLPNSCSLPYHYVNIKAIKRGFLQSDPFSHHD	299
Qy	301	HAKWCISQKTKNRTWICIGDLNRSFHOAPSSGGFCITQNWQIYQAFQGLVLYESCK	356
Db	300	HSKWCYSTKDSQARTWICIGDLNRSFHOAPSSGGFCISKNNRYIQSFEDRLVSHYASC	355

RESULT 2

JE0206

deoxyribonuclease II (EC 3.1.22.1) - human

C:Species: Homo sapiens (man)

C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Dec-2002
 C/Accession: J02006
 R/Shiokawa, D.; Tanuma, S.
 Biochem. Biophys. Res. Commun. 247, 864-869, 1998
 A/Title: Cloning of cDNAs encoding porcine and human DNase II.
 A/Reference number: J02005; MUID:98321218; PMID:9647784
 A/Accession: J02006
 A/Molecule type: mRNA
 A/Residues: 1-360 <SHI>
 A/Cross-references: GB:AF060222; NID:g3309154; PIDN:AA39852.1; PID:g3309155
 C/Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides
 C/Genetics:
 A/Map position: 19p13.2
 C/Superfamily: deoxyribonuclease II
 C/Keywords: hydrolase

Query Match 31.7%; Score 613; DB 2; Length 360;
 Best Local Similarity 36.9%; Pred. No. 1e-44;
 Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFGLVGAATISCRNEEGKAVDFTFYKLPK-RONKESGETGLLEYLYLDTSTRWRK 71
 DB 4 LLLAALLCVPAGALTCTGDSGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSGGWRD 63
 QY 72 SEQLMNDTKSVLGRITLQOLYEAYAKSNNTAYLIYNDGVKPK--VNSRYKGYHTKGLLLW 129
 DB 64 GRALINSPGAVGRSLQPLYR---SNTSQAFLYNDPQPQSKAODSSMRGHTKGVLLL 120
 QY 130 NRVOGFWLHISIPQF--PPIPEGDYDPTGRRNGSGGICITTFKYNQYEAIDSOLLVCNPN 188
 DB 121 DHGGFWLHVSFNPFPASSAAAYSWPHSACTYGTLLCVSPFPFAQSKMGKQLTYTPW 180
 QY 189 VYSCSIPATPHOELHMPQLCTRASSSEIPGRLLTTLSQAQGGKFTLHFAKSGSFLDDIFA 248
 DB 181 VINYQLEGIFAQEFDPDLENVVKGHVVSQEPWNSITLTSQAQAVFQSFKFKGDDLYS 240
 QY 249 AWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
 DB 241 GMLAALGTNLQVQFVHKTGVLPSNCSDIQWVNVNQAIFPPGAPGSPFNSTEDHSKWCV 300
 QY 307 SQKGTNRWTCIGDLNRSPhQAFRSGGFICTIONWQIYQAFQGLVLYYESC 356
 DB 301 SPKGP---WTCVGMNRRNQEGEORGGTLCALPALWKAFLVKNYQPC 347

RESULT 3
 J02005
 deoxyribonuclease II (EC 3.1.22.1) - pig
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Dec-2002
 C/Accession: J02005
 R/Shiokawa, D.; Tanuma, S.
 Biochem. Biophys. Res. Commun. 247, 864-869, 1998
 A/Title: Cloning of cDNAs encoding porcine and human DNase II.
 A/Reference number: J02005; MUID:98321218; PMID:9647784
 A/Accession: J02005
 A/Molecule type: mRNA
 A/Residues: 1-364 <SHI>
 A/Cross-references: GB:AF060221; NID:g3309152; PIDN:AA39263.1; PID:g3309153
 C/Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides
 C/Superfamily: deoxyribonuclease II
 C/Keywords: hydrolase

Query Match 30.7%; Score 593; DB 2; Length 364;
 Best Local Similarity 35.2%; Pred. No. 5.4e-43;
 Matches 126; Conservative 66; Mismatches 152; Indels 14; Gaps 7;

QY 6 LRTSFAILLGLFGLVGAATISCRNEEGKAVDFTFYKLPKRONK-RSGETGLLEYLYLDS 64
 DB 1 MATLSPLLALLLW--PVGITCTGDSGQPDVDFVYKLPALHSPGPDVAQSGLEYKYLDE 59
 QY 65 TTRSWRKSEQLMNDTKSVLGRITLQOLYEAYAKSNNTAYLIYNDGVKPK--FVNSRYKGYH 122

DB 60 BSGWRDAGAGSINGSTGALGRSLPLYR---NTSQAFLYNDPQPQPKYRSQHSNRGH 115
 QY 123 TKGILLNWRVGFGLHISIPQF--PPIPEGDYDPTGRRNGSGGICITTFKYNQYEAIDSQ 181
 DB 116 TKGVILLDQEGFGLHISVFPNPPSSAAATWSPSARTYGTGLICVSPFLTQFLNISRQ 175
 QY 182 LLVCNPNVYSCSIPATPHOELHMPQLCTRASSSEIPGRLLTTLSQAQGGKFTLHFAKSDS 241
 DB 176 LTYTPVNDYDKLEGDFAEKFPYLEEVYKHVLPQEPWNSVTLTSKAGASQSFACGN 235
 QY 242 FLDDIFAAMWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQ 299
 DB 236 FGDDLYSGWLAELGNLQVQFQWQSGAGILPNSCSGVQHVLDVDTQIAPGAPGNFATE 295
 QY 300 DHAKWCIQKGTKNRWTCTGDLNRSPhQAFRSGGFICTIONWQIYQAFQGLVLYYESCK 357
 DB 296 DSKWCVA---PERPWTGVMNRRNQEGEORGGTLCALPALWKAFLVKNYQPC 350

RESULT 4

T45071
 hypothetical protein R31240.2 [imported] - human
 C/Species: Homo sapiens (man)
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Dec-2002
 C/Accession: T45071
 R/Lamerdin, J.; McCreedy, P.; Stilwagen, S.; Ramirez, M.; Carrano, A.
 submitted to the EMBL Data Library, November 1996
 A/Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region
 A/Reference number: Z22906
 A/Accession: T45071
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-348 <LAM>
 A/Cross-references: EMBL:AD000092; PIDN:AAB51172.1
 A/Experimental source: cell line 5H12-B; fibroblast
 C/Genetics:
 A/Map position: 19p13.2
 A/Introns: 29/2; 89/3; 116/1; 159/1; 225/1
 C/Superfamily: deoxyribonuclease II

Query Match 30.0%; Score 580; DB 2; Length 348;
 Best Local Similarity 36.3%; Pred. No. 6.6e-42;
 Matches 127; Conservative 56; Mismatches 143; Indels 24; Gaps 7;

QY 13 LFLGLFGLVGAATISCRNEEGKAVDFTFYKLPK-RONKESGETGLLEYLYLDTSTRWRK 71
 DB 4 LLLAALLCVPAGALTCTGDSGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSGGWRD 63
 QY 72 SEQLMNDTKSVLGRITLQOLYEAYAKSNNTAYLIYNDGVKPK--VNSRYKGYHTKGLLLW 129
 DB 64 GRALINSPGAVGRSLQPLYR---SNTSQAFLYNDPQPQSKAODSSMRGHTKGVLLL 120
 QY 130 NRVOGFWLHISIPQF--PPIPEGDYDPTGRRNGSGGICITTFKYNQYEAIDSOLLVCNPN 188
 DB 121 DHGGFWLHVSFNPFPASSAAAYSWPHSACTYGTLLC-----KQTTYTPW 168
 QY 189 VYSCSIPATPHOELHMPQLCTRASSSEIPGRLLTTLSQAQGGKFTLHFAKSDSFLDDIFA 248
 DB 169 VINYQLEGIFAQEFDPDLENVVKGHVVSQEPWNSITLTSQAQAVFQSFKFKGDDLYS 228
 QY 249 AWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
 DB 229 GMLAALGTNLQVQFVHKTGVLPSNCSDIQWVNVNQAIFPPGAPGSPFNSTEDHSKWCV 288
 QY 307 SQKGTNRWTCIGDLNRSPhQAFRSGGFICTIONWQIYQAFQGLVLYYESC 356
 DB 289 SPKGP---WTCVGMNRRNQEGEORGGTLCALPALWKAFLVKNYQPC 335

RESULT 5

T19038
 hypothetical protein C07B5.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Dec-2002

C/Accession: T19038
R/Kershaw, J.
submitted to the EMBL Data Library, October 1994
A/Reference number: Z19064
A/Accession: T19038
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-375 <WIL>
A/Cross-references: EMBL:Z46266; PIDN:CAA86412.1; GSPDB:GN00028; CESP:C07B5.5
A/Experimental source: clone C07B5
C/Genetics:
A/Gene: CESP:C07B5.5
A/Map position: X
A/Introns: 35/2; 82/3; 108/3; 266/1; 308/3; 343/3
C/Superfamily: deoxyribonuclease II

Query Match 25.4%; Score 490.5; DB 2; Length 375;
Best Local Similarity 32.2%; Pred. No. 3.4e-34;
Matches 119; Conservative 65; Mismatches 147; Indels 39; Gaps 12;

```
QY 6 LRTSFALLFLGLFGLVGA-ATISCRNEEGKAVDWTFYKLPKRQKESGET-----GLEY 59
Db 3 LSPRAVLFL-LLGVSQTYAFAFCKDQSGNDVDWFAVYKMP--IEKDDGVTGLAGGVAV 59

QY 60 LYLDSITR-SWRKSEQLMNDTKSVLGRITLQOLY-----EAVASKSNNTAY 103
Db 60 YYVDVKKGTLPSTAKTLDNDQALAYTLQYYDKQNDKTFHVMVNDPEWGSKSTGK 119

QY 104 L-----IYNDGVKPVNYSRKYGHTKGLLNNRVQGFLLHSIPQFPPIPEGYDYP 156
Db 120 LEEILSNRVSYNTHEDDSTAFGHTKGTTFDGTSGVLVHVSPLFPN-PTK-YEYVP 177

QY 157 TGRNGQSGICITKYNQYEALDSQLVCNPNVYSCSIPATFHQELIHPQLCTRASSE 216
Db 178 SGHDYGQTMCLWTKFYAQLKSGTQLFNFRNIYSSNLPFNMAADNADLAKAIAGYQKG 237

QY 217 IPGRLLTLTQSAQOKFLHFAKSDSFLDDIEAANMAQRLKTHLLTETWQRKQLPSNCS 276
Db 238 QPFSVLETFNAGYSFNFPAKSEFNADLTDLVAPLTKDLDVVEW-RRGSEIFLDC 296

QY 277 LPYHYNIKAIKLSHRHSYFSSYQDHAKWCISQKGTNRKWTCTGIDINRSPHQAFRSGGFC 336
Db 297 LTHANDALS IHVGSSTAFSYTKDHSKWAHSADMTK-PWVCIGDINRMTSQYVRGGTTC 355

QY 337 TQN---WQIY 343
Db 356 ISSSFLWKAY 365
```

RESULT 6
S40996
hypothetical protein K04H4.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 03-May-1994 #sequence_revision 02-Aug-1994 #text_change 09-Sep-1997
C/Accession: S40996
R/Ainscough, R.
submitted to the EMBL Data Library, October 1993
A/Reference number: S40991
A/Accession: S40996
A/Molecule type: DNA
A/Residues: 1-516 <AIN>
A/Cross-references: EMBL:Z27078; NID:9414627; PID:9414633
C/Genetics:
A/Introns: 56/2; 126/1; 152/1; 181/2; 221/2; 255/3; 283/3; 360/1; 396/3; 426/1; 466/3

Query Match 21.4%; Score 413; DB 2; Length 516;
Best Local Similarity 29.5%; Pred. No. 2.2e-27;
Matches 103; Conservative 61; Mismatches 137; Indels 48; Gaps 10;
QY 17 LFGVLG-AATISCRNEEGK-----AVDW-----FTFYKLPKR 47
Db 171 MFDIIGKSGTIQCNIETTHKMLNCLIALNLSKFKVQSDMQHFFLLFVYKLPKL 230

```
QY 48 QNKESG---ETCLEVLYLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAAYAKSNNTAYL 104
Db 231 WNHDPNVPIISNGTGLFYDNNKNWKLMPQGDVENVNAVYTLQOYIN---SNMNTFSY 287

QY 105 IYNDGVKPVNYSRKYGHTKGLLNNRVQGFLLHSIPQFPPIPEGYDYPPTGRNGOS 164
Db 288 WYNDWFPDSTWNSGSHAKGVTFDQYTGFWLHSIPKFP--SKDMFRFFPNAHYGQM 345

QY 165 GICITFKYNQYEALDSQLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPQRLTT 224
Db 346 GICISYNTVSLATIAQQLFYNTFTYQFNLPQSFANQFPVLSQLKNKEYNKSPPILTSTKV 405

QY 225 LQSAQOKFLHFAKSDSFLDDIEAANMAQRLKTHLLTETWQRK---RELPSNCSLPYHV 281
Db 406 LKSLGGQHFRHFAKTGEWGLYSDFVGTLLKSSIKVETWNHQSGDEYNLPSVCD-PNHV 464

QY 282 YNKAIKLSRHSY---FSSYQDHAKWCISQKGTNR-----WTCIGDLNR 323
Db 465 QSTWSAKYIRLPYALDISYEDHSKFFVAYSESSKPPIPYVCIGDINR 513
```

RESULT 7

S44793

F09G8.2 protein - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001

C/Accession: S44793

R/Anderson, K.

submitted to the EMBL Data Library, February 1993

A/Description: Sequence of the C. elegans cosmid F09G8.

A/Reference number: S44796

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-238 <AND>

A/Cross-references: EMBL:L11247; NID:9156280; PID:9156283

C/Genetics:

A/Introns: 32/2; 171/3

Query Match

Best Local Similarity 16.5%; Score 319; DB 2; Length 238;

Matches 76; Conservative 41; Mismatches 106; Indels 20; Gaps 10;

```
QY 6 LRTSFALLFLGLFGLVGA-ATISCRNEEGKAVDWTFYKLPKRQKESGETGLEYLYLDS 65
Db 1 MRLYFVLIFSVPF-TTGNKGIQCKNMEGKSVDFWVVKLPKLSG--AGTSGKEFYFDAE 57
```

```
QY 66 TRSWRKSEQLMNDTKSVLGRITLQOLYEAAYAKSNNTAYLIYNDGVKPVNYSRKY-GHTK 124
Db 58 SSDWTRGNDI-NDPNVAVGATVSQVYS--ADKSN-N-FWFMYS-D--DDPIKSAADSYRGHAK 111
```

```
QY 125 GLILNNRVQGFLLHSIPQFPPIPEGYDYPPTGRNGQSGICITKYNQYEALDSQLLV 184
Db 112 GVSLFDDTTGFWLHSVPNFPPI--KFSFDPNTAKYIQGSFFCASMEVQHUTELAEHWKY 169
```

```
QY 185 CNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGR-----LFTLQSAQOKFLHFAK 239
Db 170 IQATPYIINIPEKY---ATRFPTLKNVEAKQSLPRSATQFWISKPIKTQGVTLMAVAKH 226
```

```
QY 240 DSF 242
Db 227 KKF 229
```

RESULT 8

A44164

secreted glycoprotein, 43 K - trichina

C/Species: Trichinella spiralis (trichina)

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Sep-1997

C/Accession: A44164; A60630; S27862

R/Vassilatis, D.K.; Despoimier, D.; Misek, D.E.; Polvere, R.I.; Gold, A.M.; Van der Ploeg

J. Biol. Chem. 267, 18459-18465, 1992

A;Title: Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode *Trichostrongylus axei*
A;Reference number: A44164; MUID:92406752; PMID:1382055
A;Accession: A44164
A;Molecule type: mRNA
A;Residues: 1-344 <VAS>
A;Cross-references: EMBL:M95499; NID:g162534; PTD:g162535
A;Note: sequence extracted from NCBI backbone (NCBIN:113310, NCBIPI:113321)
A;Note: part of this sequence was confirmed by protein sequencing
E;Gold, A.M.; Despommier, D.D.; Buck, S.W.
Mol. Biochem. Parasitol. 41, 187-196, 1990
A;Title: Partial characterization of two antigens secreted by L1 larvae of *Trichinella spiralis*
A;Reference number: A60630; MUID:90377287; PMID:2398916
A;Accession: A60630
A;Molecule type: protein
A;Residues: 23, 'X', 25-62, 'YGSF', 66-80, 'T', 115-116, 'X', 118-132, 249-278 <GOL>
C;Keywords: glycoprotein

Query Match	6.1%;	Score 117;	DB 2;	Length 344;
Best Local Similarity	20.7%;	Pred. No. 0.029;		
Matches	74;	Conservative	61;	Mismatches 142;
			Indels	80;
			Gaps	20;
Qy	13	LFLGLFGV-----LGAATISCRNEGKAVDWFYKLPKRONKESGCTG-LELYILDS	64	
Db	5	IFLSAFWVILHNCLQTHAANCTGTATDD--TENELLEK-----PVGLLAKIISP	53	
Qy	65	TTRSWRKSQBLMN-DTKSVLGRTLQOQYEAAYSASNNTAYLIYNDGVKPVNYSKYGHT	123	
Db	54	ANAGWANDGANMNTDSCHAL---VQTLAEWMGPILDDMTALGYSNTPPKSTITSQTTSS-S	109	
Qy	124	KGLLLNWR---VQGFLLIHSIPQ-PPPTPEEGYDYPPTGRRNGQSGICITFKKNQYEADS	180	
Db	110	KGILMFGNETTDDGFWLLIHTTERAPP--NSVAWSWPSTSEGHMALCALSISEDNVPLI--	165	
Qy	181	QLLVCPNVVYSCIPATFHOB-LIHMPLQCTRAS-----SSEIIPGRL-----	221	
Db	166	-----VPALQYQEVWLYFGVSSSEKATEFADLTSLIDGSIPTTTPPLWNQOT	212	
Qy	222	LTTTLQSAQGKFLHFAKSDSFLDDIFAAWMAQLKTHLLTETWQRKQELPNSCLPYHV	281	
Db	213	IITLNSALST--VVYSKTSRLEMYGSLAKVMVNM--RIWAVTDNTLQTTG--GKI	266	
Qy	282	YNIKATKLRSHTSVSSYQDHAQKWCISOKGTQNRWTCIGDLNRSPhQAFSGGFIQTQ	338	
Db	267	GFVKVVK-SPVTDIGTQNDRSK-----DKSONAVIDD---KPVFCFTTNGSYTKQ	312	

RESULT 9
 H71525
 Probable outer membrane leader peptide (omp) CT351 - Chlamydia trachomatis (serotype D, H71525)
 C:Species: Chlamydia trachomatis
 C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 04-Feb-2000
 C:Accession: H71525
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, R.
 Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
 A:Reference number: A71570; MUID:93000809; PMID:9784136
 A:Accession: H71525
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-697 <ARN>
 A:Cross-references: GB:AE001308; GB:AE001273; NID:G3328766; PIDN:AA67946.1; PID:G3328777
 A:Experimental source: serotype D, strain UW-3/Cx
 C:Genetics:
 A:Gene: CT351

```

Query Match      5.0%; Score 97.5; DB 2; Length 697;
Best Local Similarity 24.9%; Pred.No. 3.4;
Matches 62; Conservative 25; Mismatches 109; Indels 53; Gaps 12;

QY 114 VNYSRK-----YGTGKGLLWN-RVQGF-WLIH-----SIPOFPPIPEGYDVPPTG 158
      |||||
Dh 85 VNVRGKILI CDYLEVYEYTDSCLLNRCGLSPWFIIGASTIIPSSIIIHKGIVSTSEG 144
      |||||

```

Qy	159	RNCGSGICITFKNQYEADISQLVCNPVYSCIPATEHQEL-----THMPOLCTRA	211
Dd	145	---PQRHICLSGDLYKYSS-DSVLMSGPSRLSICNTPTVLLLPQISIMPEWPKPIPTFG	200
Qy	213	SSSEIPGRLLTTLQSAQGOKFLHFPAKSDSFDDIFAAWMAQRLKTHLLTTETWQRKEQLP	272
Dd	201	GSGFGLSYLGVSYSPTSCK--HCSTTLFLDGFKKHGIG-----LGYNRFSSQENP	250
Qy	273	SNCSLPYHYVNIIKAILSRHSFYSSQDHAKWCISQKTGNRWTCIGDLNRSPHOAFRSG	332
Dd	251	SN-----AINIK-----SYIAH-RLAIDSGAKDYRLHGDFTFSKERAHLAG	292
Qy	333	GFICTIONWQ	341
Dd	293	EFHLSDGWE	301

RESULT 10

C81542 conserved hypothetical protein CP0756 [imported] - Chlamydophila pneumoniae (s
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: C81542
R:Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.
, C.C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.;
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: C81542
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-704 <RES>
A:Cross-references: GB:A8002234; GB:AE002161; NID:g7189667; PIDN:AAF38559.1; P
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0756

[illegible]

RESULT 11
S63059
hypothetical protein YNL118c - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein N1917
 C:Species: Saccharomyces cerevisiae
 C:Date: 27-Apr-1996 #sequence revision 03-May-1996 #text_change 26-May-2000
 C:Accession: S63059; S59701; S67340
 R:De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63047
 A:Accession: S63059
 A:Molecule type: DNA
 A:Residues: 1-970 <DEA>
 A:Cross-references: ENBL:771394; NID:gl302044; PID:e239800; PID:gl302045; MIPS:YNL118C
 A:Experimental source: strain S288C
 R:Tzagoloff, A.A.
 A:Description: Suppressor of a yeast pet mutant.
 A:Reference number: S59701
 A:Accession: S59701
 A:Molecule type: DNA
 A:Residues: 1-424, 'L', 426-970 <TZA>
 A:Cross-references: ENBL:143065; NID:g870733; PID:g870734
 A:Experimental source: strain D273-10B
 R:d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
 submitted to the ENBL Data Library, February 1996
 A:Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces ce
 A:Reference number: S67327
 A:Accession: S67340
 A:Molecule type: DNA
 A:Residues: 1-970 <DAN>
 A:Cross-references: ENBL:Z69382; NID:gl183941; PID:e221828; PID:gl183955
 C:Genetics:
 A:Gene: SGD:PSU1
 A:Cross-references: SGD:S0005062; MIPS:YNL118C
 A:Map position: 14L
 C:Superfamily: unassigned mutT domain proteins; mutT domain homology
 F:129-163/domain: mutT domain homology <MUTT>

Query Match 4.9%; Score 95; DB 2; Length 970;
 Best Local Similarity 22.5%; Pred. No. 8.6;
 Matches 64; Conservative 35; Mismatches 99; Indels 86; Gaps 15;
 QY 17 LFGVLGATI-----SCNEEGKAVDWTFFKLPKQNKESGEGTGLVLYLDSTTR---S 68
 Db 179 IFLSGVSEVFNFQVNEIDK-IEWDFDKISKITWK-----SNIKYLLNSNMRLPSM 233
 QY 69 W-RKSEQQLMND-----TKSVLGRT-----LQLYEAVASKSNNTAVLY 106
 Db 234 WLRQRQLKNEQDLKSAEQLKLLGITKEQIDPDGRLNMLLHTAVQANSNNA--VS 291
 QY 107 NDGVKPK-----VNSRYKYGHTKGLL----- 127
 Db 292 NGQVPSQELQHLKESQGEHNNQKQSSFSQQPSIFPSLSBPANNKNVPTWEMA 351
 QY 128 --LWNRVQGFWLHISIQFPPIPEGVDYPTGRRNGSGGICITFKKNQYEAIDSQLLVC 185
 Db 352 NVFMSNPQLFATMGQP-FAPFP---FMLPLTNNSNANPIPTVPFN-FNAPPNPMAFG 406
 QY 186 NPVNYSCSIIPATFHQELIHPOLCTRAS--SSSEIPGRLLITLOS 227
 Db 407 VPNNHNLGSPAV-SQPSLPAPLPRDSGYSSSPGQLLDILNS 449

RESULT 12
 B71621
 probable membrane associated protein PFB0190C - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C:Accession: B71621
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 ; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: B71621

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2295 <GAR>
 A:Cross-references: GB:AE001379; GB:AE001362; NID:G3845118; PIDN:AACT1827.1; PID:G384511
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0190C

Query Match 4.9%; Score 94.5; DB 2; Length 2295;
 Best Local Similarity 24.9%; Pred. No. 29;
 Matches 44; Conservative 25; Mismatches 79; Indels 29; Gaps 8;
 QY 7 RTSFALLFLGLFGLVGAATISCNEEGKAVDWTFFKLPKQNKESGEGTGLVLYLDSTT 66
 Db 1089 RNSYALAAALGEIYYLGNESIGIERDEIKA---FEFWKKAADQCGDTTSALSTGYAVLYDEYK 1145
 QY 67 RSWRKSEQLMN-DTKSVLGRTIQQLYEAYASKSNNTAVLYINDGVKPKVNYSRKY----- 120
 Db 1146 KFLKKEELVKMNDREDIL--TWIHLNSTKDKKNVLEMFQESSSEKKQKKKKKKKKQD 1203
 QY 121 GHTKGLLLNWRVQGFWLHISI-----POFPPIPE-----EGYDYPPTGRRNGQS 164
 Db 1204 GNTDG-----DRVDD-KIVQNVGNVFOQSYGNVDSEMGNGSIDGFSMPSPSGLIANNVS 1255

RESULT 13

DEFP36

Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) - fruit fly (Drosophila melanogaster)
 N:Alternate names: D-glucose-6-phosphate:NADP+ oxidoreductase
 C:Species: Drosophila melanogaster
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 11-Jun-1999
 C:Accession: J02072; A38006
 R:Pouts, D.; Ganguly, R.; Gutierrez, A.G.; Lucchesi, J.C.; Manning, J.E.
 Gene 63, 261-275, 1988
 A:Title: Nucleotide sequence of the Drosophila glucose-6-phosphate dehydrogenase gene an
 A:Reference number: J02072; MUID:88255872; PMID:2838391
 A:Accession: J02072
 A:Molecule type: genomic RNA
 A:Residues: 1-523 <FOU>
 A:Cross-references: GB:M26674; NID:g157468; PIDN:AAAS1463.1; PID:g157470
 A:Note: the authors translated the codon AAG for residue 214 as Asn
 C:Comment: This is the first and dominant regulatory enzyme in the hexose monophosphate s
 C:Genetics:
 A:Gene: FlyBase:Zw
 A:Cross-references: FlyBase:FBgn0004057
 A:Introns: 6/2; 94/2; 166/2
 C:Superfamily: glucose-6-phosphate dehydrogenase
 C:Keywords: oxidoreductase; pentose phosphate pathway
 F:209/Active site: Lys #status predicted

Query Match 4.9%; Score 94; DB 1; Length 523;
 Best Local Similarity 18.3%; Pred. No. 4.7;
 Matches 50; Conservative 43; Mismatches 86; Indels 94; Gaps 12;

QY 101 TAVLYINDG-VPKFVNYSRKYGHTKGLLLNWRVQGFWLHISIQFPPIPEGVDYPP--- 156
 Db 55 TLWLWYRDDLLPKPTKFC---GYARSMITVDSIKE---QCLPYMKVQPEHQKYEFWA 107
 QY 157 -----TGRNGQSGICITFKKNQYEAIDSQLLV-----CNPNVYSCSIIPATFHQELIHM 205
 Db 108 LNEVYSGRYDGRGTC-----FELLNQLEIWMENKANRIFYLALPSPVPEVTVNI 158
 QY 206 POLCTR-----ASSSEIPGRLLITLOSQGGKFLHFAKSDSLDIFDIPA 248
 Db 159 KQLCMVCGWNRVIEKPFGRDASSAGLSHDLAGLQEDQLYRIDYLGKE----- 210
 QY 249 AWAQRLKT-----HLITETWQRKQELPNSCSLPYHVNYKAIKLS-----RHSY 294
 Db 211 --MVQKLMTIRFGNKILSSITWNEE-----NIASVLITFKPFTQGRGGY 253

QY 295 FSSY-----QDHAKWCISQGTGKRWTCIGD 320
 Db 254 FDEFGIIRDVMQNHLLQILSLVAMEKPVSCHPD 286

RESULT 14
 E81286
 probable polysaccharide modification protein Cj1413c [imported] - Campylobacter jejuni
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: E81286
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, J.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: E81286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73837.1; PID:G696884
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 C:Superfamily: kpsD protein

Query Match 4.8%; Score 93.5; DB 2; Length 394;
 Best Local Similarity 20.3%; Pred. No. 3.6;
 Matches 74; Conservative 39; Mismatches 100; Indels 151; Gaps 17;

QY 43 KLPKQNK--ESGETGLEVLYLDSTTRSWRKSEQLMNDTKSVLGRVLQOLYEAYASKNN 100
 Db 33 KVRKQNTKVKLFNFGDFFFPVSGTRC--KQDE-----KOLENFRDFFQNKKI 80

QY 101 TAYLIYNDGVKPVNYSRKYGHTK-----GLLLNWRVQGFLLIHSIPQPPPIPEEGY 152
 Db 81 DAILMYND-----CRILIAKAIKVAKELGIEIW-----IFEEGY 114

QY 153 DYPPTGRRNQSGICITFKYNOVEADSQLVCNPNVYSCSIPATPHQLIHPQLCTRA 212
 Db 115 LRP-----YCTLEKQGVNANS--LPRDKNFYLSQ-----NIFTKE 149

QY 213 SSSETPGRLTLTQSAGQOKFLHFAKSDSLDDIFA-----AW 250
 Db 150 SIKEIPG-----GFKWAF--DAFLVWLFAPILAPPFNKLIHRTLYPPFFLFW 196

QY 251 MAQRKTHLLTETWQRKQ-----ELPSNCSLPYHYVYNIAKILSRHSYFS 296
 Db 197 FRSLYKYLKITEKLNKYNLEKKYFLAILQVYSDTQIKYH-YK-KSIEHFIEITIL 254

QY 297 SYODHA-----KWCISQGTGKRWTCIGDLNRSFHOAFRSGGFICTQNWLIYQAFQ 348
 Db 255 SPANHARAKSYLVFKHHPMDRGYKNYFKLINDLSRKYH-----VEG 295

QY 349 LVLY 352
 Db 296 RVLY 299

RESULT 15
 F82873
 hypothetical protein UUS78 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
 C:Accession: F82873
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A:Reference number: A82870
 A:Accession: F82873
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-526 <GLA>
 A:Cross-references: GB:AE002156; GB:AF222894; NID:G6899580; PIDN:AAF3092.1; GSPDB:GN001
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: UUS78
 A:Genetic code: SGC3
 C:Superfamily: Ureaplasma urealyticum hypothetical protein UUS78

Query Match 4.8%; Score 93; DB 2; Length 526;
 Best Local Similarity 19.4%; Pred. No. 5.7; Indels 136; Gaps 18;
 Matches 84; Conservative 58; Mismatches 155;

QY 4 RLLRTSFALLFTGLFGVLGAATISCRN-----EERGK--AVDM 38
 Db 6 KIITSSALLFTGVATISIAACSKQNNQKTKNFNSQEQANKIIEAILNKKEGKONFIDW 65

QY 39 FTFYKLPKRONKESGETGLEVLYLDSTTRSWRKSEQLMNDTKSVLG-----RTLOQLY 91
 Db 66 VHMWNSPFOQARKLAK-DLD-KYIDLTKSKNFKSMIDKNEYKSTFSAQIENVIRIHKDI 123

QY 92 EAYAS-----KSNNTAYLIYNDGVKPVNYSRKYGHTKGLLLNWRVQGFLLIHSIPQ 143
 Db 124 KMYANSPFWKWRQOKTALFLINYA---PIQKDSQINAPGILL---PAEYPLLYSKPD 177

QY 144 FPPPIPEEGYDYPPTGRRNG-----QSGICITFKYNOYE-----AIDSO 181
 Db 178 FEELPGFGAFTPTPKSMDAVDLDTWKSFGSLISSKGSVEKSVLASKLQNSFAKTADKV 237

QY 182 LLVCN-----PNVY-----SCSIPATF-----HQLIHPQLC 209
 Db 238 VVIYDDAIVPNTYDKNGNRNLKITOFFANWMINQNYKYLAREBFLKINOKNDLIPLMST 297

QY 210 TRASSEIPG--RLTLTQSAGQOKFLHFAKSDSLDDIFAAMMAQRKTHLLTETWQRK 267
 Db 298 VHWASYGIIGMRNLITLSKAFGMP-----XDELEHLKAQEKFKIPTLKLLTND 347

QY 268 ROELPSNCSLPYHYVYNIAKILSRHSYFSYQDHAK--WCISQGTGKRWTCI-----318
 Db 348 ELELKN-----GIQTIKTGIDSPFKYHRDQNRDDWKIATNSQVLDICIALGLKPD 398

QY 319 ----GDLNRSFHO 327
 Db 399 LLVKGELSTSVHE 411

Search completed: October 14, 2004, 00:32:23
 Job time : 31.1224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:14:19 ; Search time 24.6034 Seconds
(without alignments)
755.549 Million cell updates/sec

Title: US-10-790-589-4

Perfect score: 1933

Sequence: 1 MWRLRLTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	613	31.7	360	1	DRN2_HUMAN	O00115 homo sapien
2	599	31.0	353	1	DRN2_MOUSE	P56542 mus musculus
3	593	20.7	364	1	DRN2_PIG	O62855 sus scrofa
4	505	26.1	378	1	YMW6_CAEEL	P34508 caenorhabdi
5	490.5	25.4	375	1	YK05_CAEEL	O17778 caenorhabdi
6	472.5	24.4	367	1	YLS2_CAEEL	P34387 caenorhabdi
7	459	23.7	276	1	DRN2_BOVIN	P56541 bos taurus
8	95	4.9	704	1	NEUL_HUMAN	Q9BYT8 homo sapien
9	95	4.9	970	1	PSU1_YEAST	P53550 saccharomyc
10	93.5	4.8	781	1	APB2_SULTO	Q974N6 sulfolobus
11	93	4.8	6669	1	NEBU_HUMAN	P20929 homo sapien
12	92	4.8	524	1	GSPD_DROME	P12645 drosophila
13	92	4.8	1409	1	HAP1_HAETN	P44596 haemophilus
14	91.5	4.7	305	1	Y348_MYCPN	P52555 mycoplasma
15	91	4.7	1025	1	ITR8_HUMAN	P53708 homo sapien
16	90.5	4.7	521	1	PGS1_SAPCS	P79001 saccharomyc
17	89	4.6	557	1	CNE6_HUMAN	O95741 homo sapien
18	89	4.6	557	1	CNE6_MOUSE	O92140 mus musculus
19	88.5	4.6	1456	1	MANR_HUMAN	P22897 mus sapien
20	88	4.6	424	1	VK04_VACCV	P18377 vaccinia vi
21	88	4.6	424	1	CUI8_HUMAN	Q9NVD3 homo sapien
22	88	4.6	641	1	TERL_LAMBD	P03708 bacterioph
23	87.5	4.5	358	1	Y4EE_RHLSN	P55428 rhizobium s
24	87	4.5	685	1	SC31_HUMAN	OQ7837 homo sapien
25	86	4.4	704	1	NEUL_RABIT	P42675 cryptocolag
26	85.5	4.4	508	1	TRPE_SYNY3	P20170 synecocyst
27	85.5	4.4	521	1	PGS1_YEAST	P25578 saccharomyc
28	85.5	4.4	587	1	SKU5_ARATH	Q9SU40 arabidopsi
29	85.5	4.4	775	1	ECEL_RAT	O9JHL3 rattus norv
30	85.5	4.4	951	1	SVV_ECOLI	P07118 escherichia
31	85.5	4.4	954	1	SVV_HAETN	P43834 haemophilus
32	85	4.4	386	1	SAVI_MOUSE	Q8VEB2 mus musculus
33	85	4.4	387	1	GSPL_ECOLI	P45763 escherichia

34	85	4.4	433	1	UROK_MOUSE	P06869 mus musculus
35	85	4.4	655	1	SCAG_MOUSE	Q9WU39 mus musculus
36	85	4.4	704	1	NEUL_PIG	O02038 sus scrofa
37	85	4.4	833	1	CNSA_RAT	O54735 rattus norv
38	85	4.4	880	1	LYTD_BACSU	P19848 bacillus su
39	84.5	4.4	399	1	GLF_MYCPN	P75499 mycoplasma
40	84.5	4.4	426	1	Y958_METJA	Q58368 methanococc
41	84.5	4.4	557	1	ALO_CANAL	O93852 candida alb
42	84	4.3	383	1	SAVI_HUMAN	Q9H4B6 homo sapien
43	84	4.3	475	1	RP54_THIFE	P24695 thriobacillu
44	84	4.3	1458	1	PA2R_RABIT	P49260 oryctolagus
45	83.5	4.3	1256	1	SA3_RAT	Q99M76 rattus norv

ALIGNMENTS

RESULT 1
DRN2_HUMAN
ID DRN2_HUMAN STANDARD; PRT; 360 AA.
AC O00115; O43910;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)
DE (Lysosomal DNase II) (R31240_2).
GN DNASE2 OR DNL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112802; PubMed=9446563;
RA Yasuda T., Takeshita H., Iida R., Nakajima T., Hosomi O.,
RA Nakashima Y., Kishi K.;
RT "Molecular cloning of the cDNA encoding human deoxyribonuclease II.";
RL J. Biol. Chem. 273:2610-2616(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363736; PubMed=10903447;
RA Krieser R.J., Eastman A.;
RT "Deoxyribonuclease II: structure and chromosomal localization of the
RT murine gene, and comparison with the genomic structure of the human
RT and three C. elegans homologs.";
RL Gene 252:155-162(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98382521; PubMed=9714827;
RA Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;
RT "Molecular cloning and characterization of human and murine DNase
RT II.";
RL Gene 215:281-289(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=98321218; PubMed=9647784;
RA Shiohawa D., Tanuma S.-I.;
RT "Cloning of cDNAs encoding porcine and human DNase II.";
RL Biochem. Biophys. Res. Commun. 247:864-869(1998).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=99123683; PubMed=9924608;
RA Yasuda T., Takeshita H., Iida R., Tsutsumi S., Nakajima T., Hosomi O.,
RA Nakashima Y., Mori S., Kishi K.;
RT "Structure and organization of the human deoxyribonuclease II (DNase
RT II) gene.";
RL Ann. Hum. Genet. 62:299-305(1998).
RN [6]
RP SEQUENCE FROM N.A.
RA Lamerdin J.B., McCready P.M., Stilwagen S., Ramirez M., Carrano A.;
RT "Characterization by genomic sequence analysis of a gene-rich 111 kb
RT region of 19p13.2 containing the human DNA repair gene, RAD23A.";
RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

[7] SEQUENCE FROM N.A.
 RN MEDLINE=22388257; PubMed=12477932;
 RP TISSUE=Muscle;
 RC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Cunnatne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny N.K., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [8]
 RN CARBOHYDRATE-LINKAGE SITE ASN-212.
 RP MEDLINE=22660472; PubMed=12754519;
 RX Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RA "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry";
 RL Nat. Biotechnol. 21:660-666 (2003).
 CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference
 for double-stranded DNA. Has a possible role in apoptosis.
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
 phosphates and 3'-phosphooligonucleotide end-products.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Glycosylated.
 CC -!- SIMILARITY: Belongs to the DNase II family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB004574; BAB28623.1; -
 CC EMBL; AF047016; AAC77366.1; -
 CC EMBL; AF045937; AAC35751.1; -
 CC EMBL; AF060222; AAC39852.1; -
 CC EMBL; AB008564; BAB55598.1; -
 CC EMBL; AB008559; BAB55598.1; JOINED.
 CC EMBL; AB008560; BAB55598.1; JOINED.
 CC EMBL; AB008561; BAB55598.1; JOINED.
 CC EMBL; AB008562; BAB55598.1; JOINED.
 CC EMBL; AB008563; BAB55598.1; JOINED.
 CC EMBL; AD000092; AAB51172.1; ALT_SEQ.
 CC EMBL; BC010419; AAH10419.1; -
 CC PIR; JE0206; JE0206.
 CC Genew; HGNC:2960; DNASE2.
 CC MIM; 126350; -
 CC GO; GO:0005764; C:lysosome; TAS.
 CC GO; GO:0004531; F:deoxyribonuclease II activity; TAS.
 CC GO; GO:0003677; F:DNA binding; TAS.
 CC GO; GO:0006259; P:DNA metabolism; TAS.
 CC InterPro; IPR004947; DNase_II.
 CC Pfam; PF02265; DNase_II_1.
 CC Hydrolyase, Endonuclease, Nuclease, Lysosome; Zymogen; Signal;
 KW Apoptosis; Glycoprotein; Polymorphism.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 106 BY SIMILARITY.
 FT CHAIN 107 360 DBOXYRIBONUCLEASE II.

FT CARBOHYD 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 314 R->L (in dbSNP:1061192).
 FT /FTID=VAR_012044.
 SQ SEQUENCE 360 AA; 39581 MW; DFIBBFBA9A9676EA CRC64;
 Query Match 31.7%; Score 613; DB 1; Length 360;
 Best Local Similarity 36.9%; Pred No. 2,1e-45;
 Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;
 QY 13 LFLGFLGVLGAATISCRNEEGKAVDMFTFYKLPK-RQNKESGETGLEIYLYDSTRWRK 71
 DB 4 LLLAALLCVPAAGALTCYDGGQVDFWVYVVPALRGSGEAAQRLQYKYLDESSGMRD 63
 QY 72 SSQLMNDYKSVLGRLOLQIYEAYASKNNATVLYNDGVKPK--VNYSRKYGHKGLILW 129
 DB 64 GRALINSPEGAVGRSLQPLR---SNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLL 120
 QY 130 NRVQGFHLHSIPQF--PIPERGYDYPPTGRNGSGGICITFYKQYEAIDQLLYCNPN 188
 DB 121 DHGGFWLHVHVPNPPPPASSAAYSWPHSACTYGTLLCVSPFPFAQFKMGKOLTYTPW 180
 QY 189 VYCSIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOKFLHFAKSDSFLDDIYA 248
 DB 181 VINYQLEGIFAEFPDLENVWVGHHVQSFPWNSITLTSQAGAVFQSFARFKSGDDLYS 240
 QY 249 AWAQRLKTHLTETWQRKQELPNSCLPYHYVNTKAIKLSRHS--YFSSYODHAKWCI 306
 DB 241 GWLAAALGNLQVQFWHKTGVLPSNCSDIQWLVNQLAFPGPAGPSPFNSTEDHSKCV 300
 QY 307 SOKGTNRKWTCTGDLNRSRPHQAFRSGGFICTQNWQIYQAFQGLVLYYESC 356
 DB 301 SPKGP---WTCVGMNMQEQRGGGTLCALPALWKAQPLVKNYQPC 347
 RESULT 2
 DRN2 MOUSE STANDARD; PRT; 353 AA.
 ID DRN2_MOUSE STANDARD; PRT; 353 AA.
 AC P56542; O55053;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)
 DE (Lysosomal DNase II).
 DE DNASE2 OR DNASE2A OR DNLE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98382521; PubMed=9714827;
 RA Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;
 RT "Molecular cloning and characterization of human and murine DNase
 RT II.";
 RL Gene 215:281-289 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20363736; PubMed=10903447;
 RA Krieser R.J., Eastman A.;
 RT "Deoxyribonuclease II: structure and chromosomal localization of the
 RT murine gene, and comparison with the genomic structure of the human
 RT and three C. elegans homologs.";
 RL Gene 252:155-162 (2000).
 CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference
 for double-stranded DNA. Has a possible role in apoptosis.
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
 phosphates and 3'-phosphooligonucleotide end-products.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: Belongs to the DNase II family.

ID_YMV6 CAPELL

AC	P34508; Q9NAQ0; Q9NCF8;	111	111	N-LINKED (GLCNAC. . .) (POTENTIAL).
DT	01-FEB-1994 (Rel. 28, Created)	364	378	ATLTVLLIATITFFK -> VDLANNRYF (in isoform a).
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DT	Hypothetical protein K04H4.6 in chromosome III precursor.			
DE	K04H4.6.			
GN	Caenorhabditis elegans.			
OS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Bristol N2;			
RC	MEDLINE=94150718; PubMed=7906398;			
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coultson A.,			
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,			
RA	Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,			
RA	Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,			
RA	Johnston L., Jones M., Kershaw J., Kirsten J., Laistner N.,			
RA	Laruelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,			
RA	Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,			
RA	Sims M., Smalton N., Smith A., Smith M., Sonnhammer E., Staden R.,			
RA	Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,			
RA	Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,			
RA	Wohlman P.			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.			
RT	elegans.;"			
RT	Nature 368:32-38(1994).			
RP	[2]			
RN	Nature 368:32-38(1994).			
RP	Revisions, and alternative splicing.			
RP	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.			
RL	Sequence of 18-369 from N.A. (ISOFORM A).			
RP	MEDLINE=20363736; PubMed=10903447;			
RX	Krieser R.J., Eastman A.			
RT	"Deoxyribonuclease II: structure and chromosomal localization of the			
RT	murine gene, and comparison with the genomic structure of the human			
RT	and three C. elegans homologs.;"			
RT	Gene 252:155-162(2000).			
CC	[-] ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=b;			
CC	Isoid=P34508-1; Sequence=Displayed;			
CC	Name=a;			
CC	Isoid=P34508-2; Sequence=VSP_001304;			
CC	Note=No experimental confirmation available;			
CC	[-] SIMILARITY: Belongs to the DNase II family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sb-sib.ch).			
DR	EMBL; Z19153; CAD45582.1; -.			
DR	EMBL; Z27078; CAD45582.1; JOINED.			
DR	EMBL; Z27078; CAA81586.2; -.			
DR	EMBL; Z19153; CAA81586.2; JOINED.			
DR	EMBL; Z27078; CAB76843.1; -.			
DR	EMBL; AF220525; AAP43008.1; -.			
DR	PIR; S40936; S40996.			
DR	WormPep; K04H4.6a; CE25043.			
DR	WormPep; K04H4.6b; CE31813.			
DR	InterPro; IPR004947; DNase II.			
DR	Pfam; PF03265; DNase II; 1.			
KW	Hypothetical protein; Hydrolase; Signal; Alternative splicing.			
FT	SIGNAL 1 17			POTENTIAL.
FT	CHAIN 18 378			HYPOTHETICAL PROTEIN K04H4.6.
FT	CARBOHYD 51 51			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 92 92			N-LINKED (GLCNAC. . .) (POTENTIAL).

AC	P34508; Q9NAQ0; Q9NCF8;	111	111	N-LINKED (GLCNAC. . .) (POTENTIAL).
DT	01-FEB-1994 (Rel. 26, Created)	364	378	ATLTVLLIATITFFK -> VDLANNRYF (in isoform a).
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DT	Hypothetical protein K04H4.6 in chromosome III precursor.			
DE	K04H4.6.			
GN	Caenorhabditis elegans.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;			
OC	Rhabditiidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Bristol N2;			
RC	MEDLINE=94150718; PubMed=7906398;			
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coultson A.,			
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,			
RA	Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,			
RA	Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,			
RA	Johnston L., Jones M., Kershaw J., Kirsten J., Laistner N.,			
RA	Laruelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,			
RA	Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,			
RA	Sims M., Smalton N., Smith A., Smith M., Sonnhammer E., Staden R.,			
RA	Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,			
RA	Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,			
RA	Wohlman P.;			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.			
RT	elegans.;"			
RT	Nature 368:32-38(1994).			
[2]				
RN	Nature 368:32-38(1994).			
RP	Revisions, and alternative splicing.			
RP	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.			
RL				
RL	SEQUENCE OF 18-369 FROM N.A. (ISOFORM A).			
RP	MEDLINE=20363736; PubMed=10903447;			
RX	Krieser R.J., Eastman A.;			
RT	"Deoxyribonuclease II: structure and chromosomal localization of the			
RT	murine gene, and comparison with the genomic structure of the human			
RT	and three C. elegans homologs.;"			
RT	Gene 252:155-162(2000).			
CC	[-] ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=b;			
CC	Isoid=P34508-1; Sequence=Displayed;			
CC	Name=a;			
CC	Isoid=P34508-2; Sequence=VSP_001304;			
CC	Note=No experimental confirmation available;			
CC	[-] SIMILARITY: Belongs to the DNase II family.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sb-sib.ch).			
CC				
DR	EMBL; Z19153; CAD45582.1; -.			
DR	EMBL; Z27078; CAD45582.1; JOINED.			
DR	EMBL; Z27078; CAA81586.2; -.			
DR	EMBL; Z19153; CAA81586.2; JOINED.			
DR	EMBL; Z27078; CAB76843.1; -.			
DR	EMBL; AF220525; AAP43008.1; -.			
DR	PIR; S40936; S40996.			
DR	WormPep; K04H4.6a; CE25043.			
DR	WormPep; K04H4.6b; CE31813.			
DR	InterPro; IPR004947; DNase II.			
DR	Pfam; PF03265; DNase II; 1.			
KW	Hypothetical protein; Hydrolase; Signal; Alternative splicing.			
FT	SIGNAL 1 17 POTENTIAL.			
FT	CHAIN 18 378 HYPOTHETICAL PROTEIN K04H4.6.			
FT	CHAIN 51 51 HYPOTHETICAL (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 92 92			

Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)

Db 536 VWDVDSLRSLSKHYKDGSPFIADDDLEKLVASRLVNTGLTLRQIVLSKVDQSLHTNTSL- 594
 Qy 279 YHVYNIKAIKLSRHSYFSYQDHAKW 305
 Db 595 -----DAASEYAKYC 604
 RESULT 9
 PSUL_YEAST
 ID PSUL_YEAST STANDARD; PRT; 970 AA.
 AC PS3550;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE mRNA decapping protein 2 (PSUL protein).
 GN DCP2 OR PSUL OR YNL18C OR N1917.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=D273-10B;
 RC Tragoloff A.A.;
 RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97245296; PubMed=9090055;
 RX de Antoni A., D'Angelo M.; Dai Pero F., Sartorello F., Pandolfo D.,
 RA Pallavicini A., Lanfranchi G., Valle G.;
 "The DNA sequence of cosmid 14-13b from chromosome XIV of
 RT Saccharomyces cerevisiae reveals an unusually high number of
 RT overlapping open reading frames.";
 RL Yeast 13:261-266 (1997).
 RN [3]
 RP INTERACTION WITH DCP1.
 RX MEDLINE=99438017; PubMed=10508173;
 RA Duncleley T., Parker R.;
 RT "The DCP2 protein is required for mRNA decapping in Saccharomycetes
 RT cerevisiae and contains a functional Mutr motif.";
 RX EMBO J. 18:5411-5422 (1999).
 CC -!- FUNCTION: REQUIRED FOR THE PRODUCTION OF ACTIVE DECAPPING ENZYME,
 CC PERHAPS IN A PROCESS REQUIRING THE HYDROLYSIS OF A PYROPHOSPHATE
 CC BOND. DECAPPING IS A THE MAJOR PATHWAY OF MRNA DEGRADATION IN
 CC YEAST. IT OCCURS THROUGH DEADENYLATION, DECAPPING AND SUBSEQUENT
 CC 5' TO 3' EXONUCLEOLYTIC DECAY OF THE TRANSCRIPT BODY.
 CC -!- SUBUNIT: INTERACTS WITH DCP1.
 CC -!- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. STRONG, TO
 CC S. POME SPAC19A8.12.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL; L43065; AAA68866.1; -
 DR EMBL; 269382; CAA93389.1; -
 DR EMBL; 271394; CAA95998.1; -
 DR PIR; S63059; S63059.
 DR Germonline; 143124; -
 DR SGD; S0005062; DCP2.
 DR GO; GO:0000932; C:cytoplasmic mRNA processing body; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0008047; P:enzyme activator activity; IDA.
 DR GO; GO:0003729; P:mRNA binding; IPI.
 DR GO; GO:0000290; P:deadenylation-dependent decapping; IDA.
 DR InterPro; IPR007722; DCP2.
 DR InterPro; IPR000086; NUDIX_hydrolase.
 DR Pfam; PF05026; DCP2.1.

DR InterPro; IPR001930; Peptidase M1.
 DR Pfam; PF01433; Peptidase M1; 1.
 DR PRINTS; PR00756; ALADIPTRASE.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Aminopeptidase; Zinc; Complete proteome.
 FT METAL 272 273 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 273 273 BY SIMILARITY.
 FT METAL 276 276 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 295 295 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 358 358 PROTON DONOR (POTENTIAL).
 SQ SEQUENCE 781 AA; 90370 MW; 58BBF8386DF623FB CRC64;
 Query Match 4.8%; Score 93.5; DB 1; Length 781;
 Best Local Similarity 21.2%; Pred. No. 3.7;
 Matches 67; Conservative 42; Mismatches 114; Indels 93; Gaps 14;
 QY 38 WFTFYKLPKRNKESGETGLVLYLDST-----RSWRKSEQLMMDTKSVLGRITQ 88
 Db 316 WGTFI-----NSETSGA-----LFRSLTTHPIEAHVTSPEIEQLFDDISVGKGASIL 365
 QY 89 QLYEAVASKNNTAYLIYNDGPKVPYNSRKYGHKGLLLNRVQ----- 133
 Db 366 RMEATVLGDED-----FRKGIQIYLN-TVKYSNATGSDFNLSLEKSGKPVSEIVKDWI 418
 QY 134 ---GFWLIHSIOPPIPEG-----YDYPTGRNGQSGICITPKYQVYE 177
 Db 419 TKDGYPVVYVNSGKINLEQERFYLKNGKNVAVKPLFLVNGRIYVLEKEKDSID 478
 QY 178 IDSQLLVCNPN-----VYSCSIPATFHOELHMPQ--LCTRASSEIPGRL-LTTLQ 226
 Db 479 IGSIDIKSIKNIDRTGFYRYVNDLSLVFNSKLSHLDKNGLFNDYFNFFLAGRVNVTYE 538
 QY 227 SAQQKFLHFAKSDSFL-----DDIPAAWAQRLKTHLTTWQKQKQELPNCISLPYH 280
 Db 539 SIAKQ----FMKDDNYLVDELSELVYLWVNRDKYKLLYEV-----LPYQ 581
 QY 281 V--YNIKAIKLSHSY 294
 Db 582 VKRFSKREKDELSRRTY 597
 RESULT 11
 NEBU HUMAN STANDARD; PRT; 6669 AA.
 AC P20929; Q15346;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nebulin.
 GN NEB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95257391; PubMed=7739042;
 RA Labeit S., Kolmerer B.;
 RA "The complete primary structure of human nebulin and its correlation
 to muscle structure";
 RT J. Mol. Biol. 248:308-315(1995).
 RN [2]
 RP PARTIAL PRELIMINARY SEQUENCE.
 RX MEDLINE=88284704; PubMed=3397062;
 RA Zeviani M., Darras B.T., Rizzuto R., Salviati G., Betto R.,
 RA Bonilla E., Miranda A.F., Du J., Samitt C., Dickson G., Walsh F.S.,
 RA Dimauro S., Francke U., Schon E.A.;
 RT "Cloning and expression of human nebulin cDNAs and assignment of the
 gene to chromosome 2q31-q32";
 RL Genomics 2:249-256(1988).
 RN [3]
 RP STRUCTURE BY NMR OF 6610-6669.
 RX MEDLINE=98179559; PubMed=9514727;
 RA Politou A.S., Millevoi S., Gautel M., Kolmerer B., Pastore A.;
 "SH3 in muscles: solution structure of the SH3 domain from nebulin";
 J. Mol. Biol. 276:189-202(1998).
 CC -!- FUNCTION: This giant muscle protein may be involved in maintaining
 the structural integrity of sarcomeres and the membrane system
 associated with the myofibrils. Bind and stabilize F-actin.
 CC -!- TISSUE SPECIFICITY: Muscle specific. Located in the thin filament
 of striated muscle.
 CC -!- DISEASE: Defects in NEB are a cause of the autosomal recessive
 form of nemaline myopathy (NEM2).
 CC -!- SIMILARITY: Contains 1 SH3 domain.
 CC -!- SIMILARITY: Contains 178 nebulin repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X83957; CAA58788.1; -;
 CC EMBL; M19668; AAS59916.1; ALT_SEQ.
 CC EMBL; M19669; AAS59917.1; ALT_SEQ.
 CC PIR; A29979; A29979.
 CC PIR; B29979; B29979.
 CC PIR; S55024; S55024.
 CC PDB; 1ARK; 28-JAN-98.
 CC PDB; 1NEB; 24-DEC-97.
 CC Genew; HGNC:7720; NEB.
 CC MIM; 161650; -;
 CC MIM; 256030; -;
 CC GO; GO:0015629; C:actin cytoskeleton; TAS.
 CC GO; GO:0030017; C:sarcomere; NAS.
 CC GO; GO:0008307; F:structural constituent of muscle; TAS.
 CC GO; GO:0030832; P:regulation of actin filament length; NAS.
 CC GO; GO:0007525; P:somatic muscle development; NAS.
 CC InterPro; IPR001452; SH3.
 CC Pfam; PF00880; Nebulin; 146.
 CC Pfam; PF00018; SH3; 1.
 CC PRINTS; PR00510; NEBULIN.
 CC PRINTS; PR00452; SH3DOMAIN.
 CC PRODOM; PD000066; SH3; 1.
 CC PROSITE; PS50002; SH3; 1.
 CC Actin-binding; Muscle protein; Cytoskeleton; Repeat; 3D-structure.
 FT REPEAT 76 107 NEBULIN 1.
 FT REPEAT 113 143 NEBULIN 2.
 FT REPEAT 148 178 NEBULIN 3.
 FT REPEAT 183 213 NEBULIN 4.
 FT REPEAT 218 248 NEBULIN 5.
 FT REPEAT 253 283 NEBULIN 6.
 FT REPEAT 289 318 NEBULIN 7.
 FT REPEAT 324 354 NEBULIN 8.
 FT REPEAT 363 393 NEBULIN 9.
 FT REPEAT 398 428 NEBULIN 10.
 FT REPEAT 434 464 NEBULIN 11.
 FT REPEAT 502 532 NEBULIN 12.
 FT REPEAT 537 567 NEBULIN 13.
 FT REPEAT 573 603 NEBULIN 14.
 FT REPEAT 611 641 NEBULIN 15.
 FT REPEAT 681 711 NEBULIN 16.
 FT REPEAT 749 779 NEBULIN 17.
 FT REPEAT 784 814 NEBULIN 18.
 FT REPEAT 820 850 NEBULIN 19.
 FT REPEAT 858 888 NEBULIN 20.
 FT REPEAT 893 923 NEBULIN 21.
 FT REPEAT 924 954 NEBULIN 22.
 FT REPEAT 959 990 NEBULIN 23.
 FT REPEAT 993 1023 NEBULIN 24.
 FT REPEAT 1028 1058 NEBULIN 25.
 FT REPEAT 1064 1094 NEBULIN 26.
 FT REPEAT 1102 1132 NEBULIN 27.

FT	REPEAT	1137	1167	NEBULIN 28.
FT	REPEAT	1168	1198	NEBULIN 29.
FT	REPEAT	1204	1234	NEBULIN 30.
FT	REPEAT	1237	1267	NEBULIN 31.
FT	REPEAT	1272	1302	NEBULIN 32.
FT	REPEAT	1308	1338	NEBULIN 33.
FT	REPEAT	1346	1376	NEBULIN 34.
FT	REPEAT	1381	1411	NEBULIN 35.
FT	REPEAT	1412	1442	NEBULIN 36.
FT	REPEAT	1448	1478	NEBULIN 37.
FT	REPEAT	1481	1511	NEBULIN 38.
FT	REPEAT	1516	1546	NEBULIN 39.
FT	REPEAT	1552	1582	NEBULIN 40.
FT	REPEAT	1590	1620	NEBULIN 41.
FT	REPEAT	1625	1655	NEBULIN 42.
FT	REPEAT	1656	1686	NEBULIN 43.
FT	REPEAT	1692	1722	NEBULIN 44.
FT	REPEAT	1725	1755	NEBULIN 45.
FT	REPEAT	1760	1790	NEBULIN 46.
FT	REPEAT	1796	1826	NEBULIN 47.
FT	REPEAT	1834	1864	NEBULIN 48.
FT	REPEAT	1869	1899	NEBULIN 49.
FT	REPEAT	1900	1930	NEBULIN 50.
FT	REPEAT	1936	1966	NEBULIN 51.
FT	REPEAT	1969	1999	NEBULIN 52.
FT	REPEAT	2004	2034	NEBULIN 53.
FT	REPEAT	2040	2070	NEBULIN 54.
FT	REPEAT	2078	2108	NEBULIN 55.
FT	REPEAT	2113	2143	NEBULIN 56.
FT	REPEAT	2144	2174	NEBULIN 57.
FT	REPEAT	2180	2210	NEBULIN 58.
FT	REPEAT	2213	2243	NEBULIN 59.
FT	REPEAT	2248	2278	NEBULIN 60.
FT	REPEAT	2284	2314	NEBULIN 61.
FT	REPEAT	2322	2352	NEBULIN 62.
FT	REPEAT	2357	2387	NEBULIN 63.
FT	REPEAT	2388	2418	NEBULIN 64.
FT	REPEAT	2423	2453	NEBULIN 65.
FT	REPEAT	2456	2486	NEBULIN 66.
FT	REPEAT	2491	2521	NEBULIN 67.
FT	REPEAT	2527	2557	NEBULIN 68.
FT	REPEAT	2565	2595	NEBULIN 69.
FT	REPEAT	2600	2630	NEBULIN 70.
FT	REPEAT	2631	2661	NEBULIN 71.
FT	REPEAT	2666	2696	NEBULIN 72.
FT	REPEAT	2699	2729	NEBULIN 73.
FT	REPEAT	2734	2764	NEBULIN 74.
FT	REPEAT	2770	2800	NEBULIN 75.
FT	REPEAT	2808	2838	NEBULIN 76.
FT	REPEAT	2843	2873	NEBULIN 77.
FT	REPEAT	2874	2904	NEBULIN 78.
FT	REPEAT	2909	2939	NEBULIN 79.
FT	REPEAT	2942	2972	NEBULIN 80.
FT	REPEAT	2977	3007	NEBULIN 81.
FT	REPEAT	3013	3043	NEBULIN 82.
FT	REPEAT	3051	3081	NEBULIN 83.
FT	REPEAT	3086	3116	NEBULIN 84.
FT	REPEAT	3117	3147	NEBULIN 85.
FT	REPEAT	3152	3182	NEBULIN 86.
FT	REPEAT	3185	3215	NEBULIN 87.
FT	REPEAT	3220	3250	NEBULIN 88.
FT	REPEAT	3256	3286	NEBULIN 89.
FT	REPEAT	3294	3324	NEBULIN 90.
FT	REPEAT	3329	3359	NEBULIN 91.
FT	REPEAT	3360	3390	NEBULIN 92.
FT	REPEAT	3395	3425	NEBULIN 93.
FT	REPEAT	3428	3458	NEBULIN 94.
FT	REPEAT	3463	3493	NEBULIN 95.
FT	REPEAT	3499	3529	NEBULIN 96.
FT	REPEAT	3537	3567	NEBULIN 97.
FT	REPEAT	3572	3602	NEBULIN 98.
FT	REPEAT	3603	3633	NEBULIN 99.
FT	REPEAT	3638	3668	NEBULIN 100.

FT	REPEAT	3671	3701	NEBULIN 101.
FT	REPEAT	3706	3736	NEBULIN 102.
FT	REPEAT	3742	3772	NEBULIN 103.
FT	REPEAT	3780	3810	NEBULIN 104.
FT	REPEAT	3815	3845	NEBULIN 105.
FT	REPEAT	3846	3876	NEBULIN 106.
FT	REPEAT	3914	3944	NEBULIN 107.
FT	REPEAT	3949	3979	NEBULIN 108.
FT	REPEAT	3984	4014	NEBULIN 109.
FT	REPEAT	4021	4052	NEBULIN 110.
FT	REPEAT	4057	4087	NEBULIN 111.
FT	REPEAT	4088	4118	NEBULIN 112.
FT	REPEAT	4123	4153	NEBULIN 113.
FT	REPEAT	4156	4186	NEBULIN 114.
FT	REPEAT	4191	4220	NEBULIN 115.
FT	REPEAT	4226	4256	NEBULIN 116.
FT	REPEAT	4264	4294	NEBULIN 117.
FT	REPEAT	4299	4329	NEBULIN 118.
FT	REPEAT	4330	4360	NEBULIN 119.
FT	REPEAT	4365	4395	NEBULIN 120.
FT	REPEAT	4400	4430	NEBULIN 121.
FT	REPEAT	4435	4465	NEBULIN 122.
FT	REPEAT	4471	4501	NEBULIN 123.
FT	REPEAT	4544	4574	NEBULIN 124.

QY	34	KAVDMFTYKLPKRONKESGTEGLYLY	4.8%; Score 93; DB 1; Length 6669;
Db	1933	KGIGWLPGLSLEAKNKKAMEIIEKKYRQHPDTIKYSTLMDSMNVLQAQNAKIMNEHL	Best Local Similarity 16.1%; Pred. No. 65;
QY	67	--RSWRKSE--QLMNDTKSVL-----GRTLQOLYEAYASKNNNTATLIYNDGVKPKV- 114	Matches 78; Conservative 49; Mismatches 142; Indels 214; Gaps 16;
Db	1993	YKQAWADKTKVHIMEDIPQIILAKANAINISDKLYKLSLESKKKGYDLPAIPKAA	
QY	115	-----NYSRKYGHTKGLLLMNRVQGF-----WLIHSIP-----QFPPIPERGYDYP 155	
Db	2053	KASRDIAADYKYKYNKYGK--GKMGVGRSLEDDEPKLVHSMQVAKMQSDREYKKNYENT 2109	
QY	156	PTGRRNGSGICITFKYNOYEADISOLLVCNPNVYSCSIPATFHOELIHMQLCTRASSS 215	
Db	2110	KTSVHTPADMLSVT-----AAKQAOANIITNT-----YKHLTHKYLILPDMNI 2153	
QY	216	EIPGRLLTTLQS-----AQGQKF-----LH 235	
Db	2154	ELT-RNMNRIGSDNEYKQDYNEWYKGLGWSFAGSLEVEKAKKATEYASDQYRHPSPNFQ 2212	
QY	236	FAKSDSFDDIPAAWMAQRLKTHLLTTWQRK----- 268	
Db	2213	FKKLTDSMDMVLAKQNAHTNNKHLTYTIDWNKDKTKIHWPDPDILQAKQNTLYSQKLY 2272	
QY	269	-----QELPNSCSLPYHVNIKAIKLSR-----HSYSSSYODHAKWC 305	
Db	2273	KLGWEEALKGYDLPVDVAISVQLAKASRDIAADYKYKQYRKLGHVGRSLQDDPKLY 2332	
QY	306	ISQKGTG-----NRWTCIG 319	
Db	2333	LSMNVAKQSEREYKQDFEKWKTKFSPFVDMGLGVLLAKKQCELVSVDYKYNLHQWTCILP 2392	
QY	320	DLN 322	
Db	2393	DQN 2395	

RESULT 12	
G6PD_DROME	STANDARD; PRT; 524 AA.
ID	G6PD_DROME
AC	P12646; Q27574; Q27872; Q27879; Q27881; Q9VWE2; Q9VWE3;
DT	01-OCT-1989 (Rel. 12, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)

DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (G6PD)
DE (Zwischenferment).
GN ZW OR G6PD OR CG12529.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88255872; PubMed=2833891;
RA Fouts D., Ganguly R., Gutierrez A.G., Lucchesi J.C., Manning J.E.;
RT "Nucleotide sequence of the Drosophila glucose-6-phosphate
RT dehydrogenase gene and comparison with the homologous human gene.";
RL Gene 63:261-275(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P23.3, F24.1, MT41, MT68, Z3, Z5, Z11, Z16, Z21, Z27, Z41, Z42,
RC Z55, Z62, Z64, and Z74;
RX MEDLINE=97070821; PubMed=8913747;
RA Banes W.F., Kirchner M., Yoon J., Biermann C.H., Wang I.N.,
RA McCartney M.A., Verrelli B.C.;
RA "Historical selection, amino acid polymorphism and lineage-specific
RT divergence at the G6pd locus in Drosophila melanogaster and D.
RT simulans.";
RT Genetics 144:1027-1041(1996).
RL [3]
RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RP STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.H., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Brattaker P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Cantor A., Chandra I.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Davies P.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalaali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM LONG).
RP STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,

George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
Rubin G.M., Celniker S.E.;
" A Drosophila full-length cDNA resource.";
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
-!- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-
1,5-lactone 6-phosphate + NADPH.
-!- PATHWAY: Pentose phosphate pathway; first step.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
IsoId=Pi2646-1; Sequence=Displayed;
Name=Short;
IsoId=Pi2646-2; Sequence=VSP 001593;
NOTE=No experimental confirmation available;
-!- SIMILARITY: Belongs to the glucose-6-phosphate dehydrogenase
family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M26674; AAA51463.1; -;
EMBL; M26673; AAA51463.1; JOINED.
EMBL; U42738; AAB02801.1; -;
EMBL; U42739; AAB02802.1; -;
EMBL; U42740; AAB02803.1; -;
EMBL; U42741; AAB02804.1; -;
EMBL; U42742; AAB02805.1; -;
EMBL; U42743; AAB02806.1; -;
EMBL; U42744; AAB02807.1; -;
EMBL; U42745; AAB02808.1; -;
EMBL; U42746; AAB02809.1; -;
EMBL; U42747; AAB02810.1; -;
EMBL; U42748; AAB02811.1; -;
EMBL; U42749; AAB02812.1; -;
EMBL; U43165; AAA99071.1; -;
EMBL; U43166; AAA99072.1; -;
EMBL; U43167; AAA99073.1; -;
EMBL; U44721; AAA99092.1; -;
EMBL; U45985; AAA99107.1; -;
EMBL; AE003512; AAF48999.1; -;
EMBL; AE003512; AAF49000.1; -;
EMBL; AY052079; AAK93503.1; -;
PIR; A47740; A47740.
PIR; J02072; DEFG6.
HSP; P11411; IDPG.
FlyBase; FBgn004057; Zw.
InterPro; IPR001282; G6PD.
Pfam; PF00479; G6PD; 1.
Pfam; PF02781; G6PD; C; 1.
PRINTS; PR00079; G6PDHGNASE.
ProDom; PD001129; G6PD; 1.
TIGRFAMs; TIGR00871; zwf; 1.
PROSITE; PS00069; G6P DEHYDROGENASE; 1.
Oxidoreductase; NADP; Glucose metabolism; Polymorphism;
Alternative splicing.
ACT SITE 209 209
FT VARSPLIC 1 7
FT Missing (in isoform Short).
FT /FTID=VSP_001593.
FT G -> C (IN STRAINS F24.1, MT32 AND MT68).
FT T -> N (IN STRAIN Z74).
FT L -> P (IN STRAINS F23.3, MT41, Z3, Z5,
FT Z11, Z16, Z21, Z27, Z42, Z55, Z64, Z74
FT AND BERKELEY).
FT QA -> AG (IN REF. 1).
FT N -> K (IN REF. 1).
FT LGV -> ARS (IN REF. 1).
FT DELRE -> AAAQ (IN REF. 1).
FT SEQUENCE 524 AA; 60431 MW; ABF81B763A82F1FD CRC64;
SQ


```

or send an email to license@isb-sib.ch).
-----
EMBL; U32710; -, NOT_ANNOTATED_CDS.
CC MEROPS; S06.006; -.
DR TIGR; H10248; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR000710; Peptidase_S6.
DR InterPro; IPR004899; Pertactin_S6.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03242; Pertactin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR TRGFAMS; TIGR01414; autotrans_bar1; 1.
CC KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal;
Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.
FT PROPEP ? 1409 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 250 250 BY SIMILARITY.
CC SQ SEQUENCE 1409 AA; 156797 NW; 63ABC893FAB84D16E CRC64;

```

```

Query Match      4.8%; Score 92; DB 1; Length 1409;
Best Local Similarity 20.4%; Pred. No. 11;
Matches         76; Conservative 49; Mismatches 124; Indels 124; Gaps 19;

QY   32 EGRAVDWFTF-YKLPRKONKESETGLE------LYL 62
     ||| :||| :||| :||| :||| :||| :||| :||| :
Db   110 EGRNPDRHFTVQIVKRNNYQAWERKHYPDGDGYHMRPHKFVTEAPVGMTTNDMGKYA 169
                                     :
QY   63 D-----STTRSWKSEQLMNDTKS-----VLGRTLQLIYEAYAKSN 100
     ||| :||| :||| :||| :||| :||| :||| :
Db   170 DRENYPERVIISGRGYWRFDKDEETNVLISSYVSAGRYLLFAGNTHQTQ-----SGNGNG 224

```

```

101 TAYLIINDGVKPVNYS-RKYGHTKG-----LLLMNRVQGFLLIHSIPOFPPIPERGYDY 154
125 TYNLSGN---VSPNHYGPLETGGSKDGSMPFIYDAKKQWMLNAVLO-----TGH 274
155 PPTGRNGOGSICINPKNOYEAISOLLVCNPNVYSCSI-PATHQOELIHPOLCTRAS 213
175 PFFGRNGQLTBREWFNEVUAVT-----PSVFORYPINGHYSEV-----S 319
214 SSEIPGRLLTTLOSAGQOKFLHFAKSDSFLDDIFAAMWAQRLKTHLLTETWORKQLPS 273
320 NNDGTGKLTLTTPSGDGSK---AKSEVGTVKLFNPSLNQTAKEHV----- 361
274 NCSLPVHVNIKAIKLSHSEYSSQDHAKCWICISOKGTYKNWTCIGDLNRSQHQAFRSGG 333
362 XAAAGVNIYQPR-MBYGKNYILGCD-----QKGKT---LTIENNINQAGGYIFEGN 408
334 FIC---TON---WQ 341
409 FVVGKKNITWQ 421

```

```

RESULT 14
Y348_MYCPN
IID_Y348_MYCPN STANDARD; PRT; 305 AA.
F75255;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical lipoprotein MG348 homolog precursor (G12_orf305).
DE MPN523 OR MP319
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_taxID=2104;
[1]
SEQUENCE FROM N.A.
STRAIN=ATCC 29342 / M129;
MEDLINE=97105885; PubMed=8948633;
Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
Herrmann R.:

```

Query Match	4.8%;	Score 92;	DB 1;	Length 524;
Best Local Similarity	18.3%;	Pred. No. 2.9;		
Matches	50;	Conservative 41;	Mismatches 88;	Indels 94;
Gaps	12;			
Qy	101	TAYLIYNDG-VPKPVNYGRKYGHTKGLLLNNRVQGFHLIHSIPFPPIPEGYDYP	---	156
Db	55	TLWMLYRDLLPKPTKFC---GYARSMLTVDISKE---QCLPYMKQPHQKQYEFWA	107	
Qy	157	---TGRNGSGGICFTKYNQYEAIDQLLV-----CNPVYSSCIPATFHOELIHM	205	
Db	108	LNKVSGRIDRTG-----PELLNQQLLEIMENKNKRIFYALPPSVFBEVTNI	158	
Qy	206	PQLCTR-----ASSSIPGRLTTTQLSQAGQKFLFAKSDSFLDDIFA	248	
Db	159	KQICMSVCGWNRVILEXPFGRDDASSQALSHLAGLFOEDLYRIDYLGKE-----	210	
Qy	249	ANMAQRUKT-----HLITETWQRKQELPNSCLPYHYNIKAIKLS-----RHSY	294	
Db	211	--MVQNLMTIRFGNKILSSWTNRE-----NIASVLIITFKPFGTQGRGGY	253	
Qy	295	FSY-----QDAKWCISOKGYKNWTCIGD	320	
Db	254	FDEGIIRDVWQNLHLLQILSLIVAMEKPSVCHPD	286	

RESULT 13	
HAP1	HAEIN
ID	HAP1_HAEIN STANDARD; PRT; 1409 AA.
AC	P4456;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Adhesion and penetration protein precursor (EC 3.4.21.-).
GN	HAP OR HI0248.
OS	Haemophilus influenzae.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC	Pasteurellaceae; Haemophilus.
OX	NCBI_TaxID=727;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Rd / KW20 / ATCC 51907;
RX	MEDLINE=95350630; PubMed=7542800;
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A. Gocayne J.D.,
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA	Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA	Venter J.C.;
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae
RL	Rd";
RL	Science 269:496-512(1995).
CC	-!- FUNCTION: PROBABLE PROTEASE; PROMOTES ADHERENCE AND INVASION BY
CC	DIRECTLY BINDING TO A HOST CELL STRUCTURE (BY SIMILARITY).
CC	-!- SUBCELLULAR LOCATION: Secreted (Potential).
CC	-!- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC	SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC	OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC	DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC	SIMILARITY).
CC	-!- SIMILARITY: Belongs to peptidase family S6.
CC	-!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON HAD TO
CC	BE SKIPPED IN POSITION 710 TO PRODUCE THIS ORF.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";
RT Nucleic Acids Res. 24:4420-4449(1996).
EL -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Potential).
CC
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AB000029; AAB95967.1; --
DR PIR; S73645; S73645.
DR InterPro; IPR000437; Prok lipoprot S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Lipoprotein; Membrane; Signal;
KW Complete proteome; Palmitate.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 305 HYPOTHETICAL LIPOPROTEIN MG348 HOMOLOG.
FT LIPID 30 30 N-palmitoyl cysteine (Potential).
FT LIPID 30 30 S-diacylglycerol cysteine (Potential).
SQ SEQUENCE 305 AA; 33536 MW; 38A143D603F4BBE3 CRC64;
Query Match 4.7%; Score 91.5; DB 1; Length 305;
Best Local Similarity 23.9%; Pred. No. 1.6;
Matches 55; Conservative 33; Mismatches 75; Indels 67; Gaps 12;
QY 7 RTSPA-LLFLGLFGLVLA-----TISCNEEGKAVDMFTFKLPKRONKE----- 51
Db 12 KVSFSTILLGSGVILSSCNIDKPNVFTLSQSVENK-VD---YSKLPK-SNKTVRNL 66
QY 52 -----SGETGLYL-YLDSTTSWRKSEQLMNDTKSVLGRTLQQLYEAY 94
Db 67 VFGTAENDGNYVLVVTETDSSQINFNGSNQAVSTENWAGD-----LGTWKQVQNY 122
QY 95 ASKSNNTAYLIYNDGVPKPNYSRKYGTGKLLMNRVQGFWLHISIPQPPPIPEGYDY 154
Db 123 STYPKGVKFLWINDIDENPKV-----WNPFRYPVIASDNLAKQTDK--DN 167
QY 155 PPTGRRNGQSGICITFKINQVHAIDSQLVCNPNVYSCSIPATPHOLIH 204
Db 168 SDKLRNDESAI-----QYR-----EIVTFIQTVYSGVNNLINQSNVH 206
RESULT 15
ID ITA8 HUMAN STANDARD; PRT; 1025 AA.
AC P53708;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-8.
GN ITGA8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95286701; PubMed=7768999;
RA Schnapp L.M., Brüss J.M., Ramos D.M., Sheppard D., Pytela R.;
RT "Sequence and tissue distribution of the human integrin alpha 8 subunit: a beta 1-associated alpha subunit expressed in smooth muscle cells.";
RL J. Cell Sci. 108:537-544(1995).
CC -!- FUNCTION: INTEGRIN ALPHA-8/BETA-1 IS A RECEPTOR FOR FIBRONECTIN AND CYTOACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A DISULFIDE BOND. ALPHA-8 ASSOCIATES WITH BETA-1.
CC

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN SMOOTH MUSCLE AND SMOOTH MUSCLE-LIKE CONTRACTILE CELLS.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; L36531; AAB93514.1; --
DR HSP; P06756; LJV2.
DR Genew; HGNC:6144; ITGA8.
DR MIM; 604063; --
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; R:cell adhesion receptor activity; NAS.
DR GO; GO:0016337; P:cell-cell adhesion; NAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 4.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 6.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium.
FT CHAIN 1 868 INTEGRIN ALPHA-8 HEAVY CHAIN (POTENTIAL).
FT CHAIN 869 1025 INTEGRIN ALPHA-8 LIGHT CHAIN (POTENTIAL).
FT DOMAIN 1 974 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 975 995 POTENTIAL.
FT DOMAIN 996 1025 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 79 FG-GAP 1.
FT REPEAT 95 160 FG-GAP 2.
FT REPEAT 161 225 FG-GAP 3.
FT REPEAT 226 279 FG-GAP 4.
FT REPEAT 280 345 FG-GAP 5.
FT REPEAT 346 405 FG-GAP 6.
FT REPEAT 409 463 FG-GAP 7.
FT CA_BIND 291 299 POTENTIAL.
FT CA_BIND 357 365 POTENTIAL.
FT CA_BIND 421 429 POTENTIAL.
FT SITE 417 419 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 58 68 BY SIMILARITY.
FT DISULFID 112 133 BY SIMILARITY.
FT DISULFID 149 162 BY SIMILARITY.
FT DISULFID 469 480 BY SIMILARITY.
FT DISULFID 486 542 BY SIMILARITY.
FT DISULFID 603 609 BY SIMILARITY.
FT DISULFID 675 688 BY SIMILARITY.
FT DISULFID 829 886 INTERCHAIN (BY SIMILARITY).
FT DISULFID 891 896 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 563 563 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 715 715 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 742 742 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 858 858 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 885 885 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 967 967 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1025 AA; 113612 MW; 58B56B2C00CF7B93 CRC64;
Query Match 4.7%; Score 91; DB 1; Length 1025;

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:33:41 ; Search time 4897.47 Seconds
(without alignments)
11221.906 Million cell updates/sec

Title: US-10-790-589-3
Perfect score: 1268
Sequence: 1 atggggaagtgctcgtg.....tctctcatgtttaccattta 1268

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hg_hum.*
- 31: em_hg_inv.*
- 32: em_hg_other.*
- 33: em_hg_mus.*
- 34: em_hg_pln.*
- 35: em_hg_rod.*
- 36: em_hg_mam.*
- 37: em_hg_vrt.*
- 38: em_sy.*
- 39: em_hg_hum.*
- 40: em_hg_mus.*
- 41: em_hg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1268	100.0	1268	6	AR201144 Sequence
2	1140	89.9	1140	9	AF274571 Homo sapi
3	1079.6	85.1	1086	6	AR432047 Sequence
4	1079.6	85.1	1086	6	BD091715 Novel deo
5	920.6	72.6	1029	9	AF333389 Homo sapi
6	881.8	53.8	1498	10	AF178974 Rattus no
7	673.8	53.1	1652	6	AR432046 Sequence
8	673.8	53.1	1652	6	BD091714 Novel deo
9	672.8	53.1	1654	10	AF128888 Mus muscu
10	672.8	53.1	1224	6	AR201143 Sequence
11	487.8	38.5	19164	9	AF334602 Homo sapi
12	487.8	38.5	164529	2	AC027606 Homo sapi
13	487.8	38.5	205034	9	AL359273 Human DNA
14	254	20.0	2337	10	AF334603S6 Mus muscu
15	254	20.0	197457	2	AC127571 Mus muscu
16	254	20.0	211003	2	AC141632 Mus muscu
17	254	20.0	255925	2	AC114618 Mus muscu
18	246.4	19.4	231698	2	AC118117 Rattus no
19	246.4	19.4	306870	2	AC098557 Rattus no
20	179	14.1	86816	2	AL136113 Homo sapi
21	160	12.6	1357	9	AB074093 Homo sapi
22	158.4	12.5	1387	9	AB074094 Pan trogl
23	156.8	12.4	1388	9	AB074095 Gorilla g
24	156.4	12.3	1388	9	AB074096 Pongo pyg
25	155.2	12.2	1390	9	AB074097 Hyllobates
26	155.2	12.2	1391	9	AB074099 Macaca mu
27	155	12.2	7470	14	F06408 Fowlpox v
28	155	12.2	288539	14	AF198100 Fowlpox v
29	153.6	12.1	1388	9	AB074098 Papio han
30	153.6	12.1	1390	9	AB074100 Macaca fa
31	148.8	11.7	1391	9	AB074101 Aotus tri
32	114	9.0	642	10	AF334603S5 Mus muscu
33	97.4	7.7	1374	10	AF334603S2 Mus muscu
34	89.2	7.0	719	10	AF334603S4 Mus muscu
35	84.8	6.7	1724	9	AK075449 Homo sapi
36	84	6.6	1593	9	AB004574 Homo sapi
37	84	6.6	1920	9	AF047016 Homo sapi
38	82.4	6.5	1083	9	BT007047 Homo sapi
39	82.4	6.5	1083	12	BT007514 Synthetic
40	82.4	6.5	1575	6	BC337903 Sequence
41	82.4	6.5	1601	9	BC010419 Homo sapi
42	82.4	6.5	1751	9	AF045937 Homo sapi
43	82.4	6.5	1825	10	AF334603S1 Mus muscu
44	82.4	6.5	1975	9	AF060222 Homo sapi
45	73	5.8	1292	4	SSJ001387 Sus scrof

ALIGNMENTS

RESULT 1	AR201144	Sequence 3 from patent US 6358723.	1268 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201144	Sequence 3 from patent US 6358723.				
DEFINITION	AR201144	Sequence 3 from patent US 6358723.				
ACCESSION	AR201144	Sequence 3 from patent US 6358723.				
VERSION	AR201144.1	GI:20252032				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1268)					
AUTHORS	Eastman, A. Richard. and Krieser, R. Joe.					
TITLE	Deoxyribonuclease II beta. proteins and cDNAs					
JOURNAL	Patent: US 6358723-A 3 19-MAR-2002;					
FEATURES	Location/Qualifiers					

[illegible]

[illegible]

Db	601	ACCTTTTCACGAGAGCTCATTCACATGCCCCAGCTGTGCACAGGCCAGCTCATCAGAG	660
Qy	694	ATTCTGGGAGGCTCTCTACACACATTCAGTCGGCCCGAGGACAAATAATCTCTCCATTTT	753
Db	661	ATTCTGGGAGGCTCTCTACACACATTCAGTCGGCCCGAGGACAAATAATCTCTCCATTTT	720
Qy	754	GCAAGTCGGATCTTTTCTTGACGACATCTTTTGACGCTGGATGGCTCAACGGCTGAAG	813
Db	721	GCAAGTCGGATCTTTTCTTGACGACATCTTTTGACGCTGGATGGCTCAACGGCTGAAG	780
Qy	814	ACACACTGTTTAACAGAACTGGCGAGGAAAGAAAGCAAGAGCTCTCTTCAAACTGCTCC	873
Db	781	ACACACTGTTTAACAGAACTGGCGAGGAAAGAAAGCAAGAGCTCTCTTCAAACTGCTCC	840
Qy	874	CTTCTTACCAGTCTACATATAAAGCAATTAATATACAGACATCTCTTATTTCAAGT	933
Db	841	CTTCTTACCAGTCTACATATAAAGCAATTAATATACAGACATCTCTTATTTCAAGT	900
Qy	934	TCCTTACAGATCAGCGCAAGTGGTGATTTTCCAAAGGGCCACCAAAATCTCTGGACA	993
Db	901	TCCTTACAGATCAGCGCAAGTGGTGATTTTCCAAAGGGCCACCAAAATCTCTGGACA	960
Qy	994	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1053
Db	961	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1020
Qy	1054	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATATATAGAAAGCTGT	1113
Db	1021	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATATATAGAAAGCTGT	1080
Qy	1114	AAGTAA 1119	
Db	1081	AAGTAA 1086	
RESULT 4			
LOCUS	BD091715	1086 bp	DNA linear
DEFINITION	Novel deoxyribonuclease, gene encoding thereof and use thereof.		PAT 27-AUG-2002
ACCESSION	BD091715		
VERSION	BD091715.1	GI:22637326	
KEYWORDS	WO 0112793-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1086)		
TITLE	Tanuma, S. and Shiohara, D.		
JOURNAL	Novel deoxyribonuclease, gene encoding thereof and use thereof		
COMMENT	Patent: WO 0112793-A 2 22-FEB-2001; SEIICHI TANUMA, DAISUKE SHIOHARA		
	OS Homo sapiens (human)		
	EN WO 0112793-A/2		
	PD 22-FEB-2001		
	PF 01-MAY-2000 WO 2000JP002893		
	PR 17-AUG-1999 JP 99P 230870		
	PI SEIICHI TANUMA, DAISUKE SHIOHARA		
	PC C12N9/22, C12N15/12, C12N5/10, C12N7/01, A61K31/195, A61K48/00		CC
	FT Key	Location/Qualifiers	
	(1)..(1083).		
FEATURES	Location/Qualifiers		
source	1..1086		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	85.1%;	Score 1079.6;	DB 6; Length 1086;
Best Local Similarity	99.6%;	Pred. No. 3.6e-291;	
Matches 1082;	Conservative	0; Mismatches 4;	Indels 0; Gaps 0;
Qy	34	ATGAACAGAAATGATGCAAGACTGCTAAGAACATCCTTTGCTTTGCTCTCTCTTGGC	93

Db	1	ATGAACAGAAATGATGCAAGACTGCTAAGAACATCCTTTGCTTTGCTCTCTCTTGGC	60
Qy	94	CTCTTTGGGGTCTGGGGGAGCAAAATTCATGAGAAATCAAGAGGAAAGCTGTG	153
Db	61	CTCTTTGGGGTCTGGGGGAGCAAAATTCATGAGAAATCAAGAGGAAAGCTGTG	120
Qy	154	GACTGGTTTACTTTTATTAAGTTTACCTAAAAGACAAAACAAGAAAGTGGAGAGCTGGG	213
Db	121	GACTGGTTTACTTTTATTAAGTTTACCTAAAAGACAAAACAAGAAAGTGGAGAGCTGGG	180
Qy	214	TTAGAGTACCTGTACCTAGACTCTCAACTAGAAAGCTGGAGGAAAGTGGAGAACTAATG	273
Db	181	TTAGAGTACCTGTACCTAGACTCTCAACTAGAAAGCTGGAGGAAAGTGGAGAACTAATG	240
Qy	274	AATGACACCAAGAGTGTGGGAAAGCAATTAACACAGCTATATGAAGCATATGCTCT	333
Db	241	AATGACACCAAGAGTGTGGGAAAGCAATTAACACAGCTATATGAAGCATATGCTCT	300
Qy	334	AAGAGTAAACAAACAGCTATCTAATATACAATGATGGAGTCCCTAAACCTGTGAATTAC	393
Db	301	AAGAGTAAACAAACAGCTATCTAATATACAATGATGGAGTCCCTAAACCTGTGAATTAC	360
Qy	394	AGTAGAAAGTATGGAACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGG	453
Db	361	AGTAGAAAGTATGGAACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGG	420
Qy	454	CTGATTCATTCCTCCTCCTCAGTTTCTCCAAATTCGGGAAGGCTATGATTTCCACCC	513
Db	421	CTGATTCATTCCTCCTCCTCAGTTTCTCCAAATTCGGGAAGGCTATGATTTCCACCC	480
Qy	514	ACAGGAGACGAAATGGAACAAAGTGGCATCTGCAATACTTTCAAGTACAACCAAGTATGAG	573
Db	481	ACAGGAGACGAAATGGAACAAAGTGGCATCTGCAATACTTTCAAGTACAACCAAGTATGAG	540
Qy	574	GCAATAGATTCTCAGTCTTGGTCTGCAACCCCAAGCTCTATAGCTCTCCATCCAGCC	633
Db	541	GCAATAGATTCTCAGTCTTGGTCTGCAACCCCAAGCTCTATAGCTCTCCATCCAGCC	600
Qy	634	ACCTTTACAGAGGCTCATTACATGCCCGAGCTGACCCAGGCGCAGCTCATCAGAG	693
Db	601	ACCTTTACAGAGGCTCATTACATGCCCGAGCTGACCCAGGCGCAGCTCATCAGAG	660
Qy	694	ATTCCTGGCAGGCTCTCACCACACTTCAGTCGGCCCGAGGACAAAATTCCTCAATTT	753
Db	661	ATTCCTGGCAGGCTCTCACCACACTTCAGTCGGCCCGAGGACAAAATTCCTCAATTT	720
Qy	754	GCAAGTCGGATCTTTTCTTGACGACATCTTTCAGCTGGATGGCTCAACGGCTGAAG	813
Db	721	GCAAGTCGGATCTTTTCTTGACGACATCTTTCAGCTGGATGGCTCAACGGCTGAAG	780
Qy	814	ACACACTGTTTAAACAGAACTGGCGAGGAAAGCAAGAGCTCTCTTCAAACTGCTCC	873
Db	781	ACACACTGTTTAAACAGAACTGGCGAGGAAAGCAAGAGCTCTCTTCAAACTGCTCC	840
Qy	874	CTTCTTACCAGTCTACATATAAAGCAATTAATATACAGACATCTCTTATTTCAAGT	933
Db	841	CTTCTTACCAGTCTACATATAAAGCAATTAATATACAGACATCTCTTATTTCAAGT	900
Qy	934	TCCTTACAGATCAGCGCAAGTGGTGATTTTCCAAAGGGCCACCAAAATCTCTGGACA	993
Db	901	TCCTTACAGATCAGCGCAAGTGGTGATTTTCCAAAGGGCCACCAAAATCTCTGGACA	960
Qy	994	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1053
Db	961	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1020
Qy	1054	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATATATAGAAAGCTGT	1113
Db	1021	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATATATAGAAAGCTGT	1080
Qy	1114	AAGTAA 1119	

Db	1081	AGTAA	1086	
RESULT 5	AF333389			
LOCUS	Homo sapiens lung-specific DLAD mRNA	1029 bp	mRNA	linear
DEFINITION	AF333389			FRI 22-NOV-2001
ACCESSION	AF333389.1	GI:17046270		complete cds.
VERSION				
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1029)			
AUTHORS	Shiokawa,D. and Tanuma,S.I.			
TITLE	Isolation and characterization of the DLAD/Dlad genes, which lie head-to-head with the genes for urate oxidase			
JOURNAL	Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)			
MEDLINE	21556924			
PUBMED	11700027			
REFERENCE	2 (bases 1 to 1029)			
AUTHORS	Shiokawa,D. and Tanuma,S.-I.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-JAN-2001) Biochemistry, Science Univ. of Tokyo, Ichigaya, Shinjuku-Ku, Tokyo 162-0826, Japan			
FEATURES	location/Qualifiers			
source	1. .1029			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
CDS	408..869			
	/note="expressed only in lung; does not include first two exons"			
	/codon_start=1			
	/product="lung-specific DLAD"			
	/protein_id="AAL34448.1"			
	/db_xref="GI:17046271"			
	/translation="MPQLCTRRSSSEIPGRLTLTQSAQOKPLHPKSDSFLDDIFLA AWAORLKHLLTETWORKQELPNSCLPYHYNIKAIKLSHRSVFSSVQDHAKWCI SOKGTGRWTCIGDLNRSRPHQAFRSGGFTCTQNWQIYQAFQGLVLYESCK"			
ORIGIN				
Query Match	72.6%;	Score	920.6;	DB 9;
Best Local Similarity	99.6%;	Pred. No.	1.3e-246;	Length 1029;
Matches	92;	Conservative	0;	Mismatches 4;
			Indels	0;
			Gaps	0;
QY	335	AGAGTAACAAACACAGCCTATCTAATATACATCATGGAGTCCCTTAACCTGTGAATTACA	394	
Db	85	AGAGTAACAAACACAGCCTATCTAATATACATCATGGAGTCCCTTAACCTGTGAATTACA	144	
QY	395	GTAGAAAGTATGGACACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGTTCTGGC	454	
Db	145	GCAGAAGTATGGACACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGTTCTGGC	204	
QY	455	TGATTTCATTCCCTCCCTCAGTTTCCTCCCAATTCGGAAGAAGCTATGATATCCACCA	514	
Db	205	TGATTTCATTCCCTCCCTCAGTTTCCTCCCAATTCGGAAGAAGCTATGATATCCACCA	264	
QY	515	CAGGGAGACAAATGGACAAAGTGCGATCTGCATAACTTTCAAGTACACCAAGTATGAGG	574	
Db	265	CAGGGAGACAAATGGACAAAGTGCGATCTGCATAACTTTCAAGTACACCAAGTATGAGG	324	
QY	575	CAATAGATTCTAGCTCTTGGTGTGCAACCCCAACGCTATATAGCTGCTCCATCCAGCCA	634	
Db	325	CAATAGATTCTAGCTCTTGGTGTGCAACCCCAACGCTATATAGCTGCTCCATCCAGCCA	384	
QY	635	CCTTTCACCAAGAGGCTCATTTCATGCCCCAGCTGTGCACAGGCGCAGCTCATCAGAGA	694	
Db	385	CCTTTCACCAAGAGGCTCATTTCATGCCCCAGCTGTGCACAGGCGCAGCTCATCAGAGA	444	
QY	695	TTCTCTGGCAGGCTCTCTCACCACTTCAGTCGCGCCACAGGACAAAAATTCCTCCATTTTG	754	

Db	445	TTCTCTGGCAGGCTCCTCACCACACTTCAGTTCGGCCCGGACAAAAATTCCTCCATTTC	508
QY	755	CAAAGTCGGATTCTTTTCTTTGACGACATCTTTGACGCTTGGATGGCTCAACGGCTGAAGA	814
Db	505	CAAAGTCGGATTCTTTTCTTTGACGACATCTTTGACGCTTGGATGGCTCAACGGCTGAAGA	564
QY	815	CACACTTGTAAACAGAAACCTGGCGACGAAAAAGACAGAGAGCTTCCTTCAAACTGCTCC	874
Db	565	CACACTTGTAAACAGAAACCTGGCGACGAAAAAGACAGAGAGCTTCCTTCAAACTGCTCC	624
QY	875	TTCTCTTACATGTCTACAAATATAAAGCAATTAATTAATCAAGACACTTTATTTTCAGTT	934
Db	625	TTCTCTTACATGTCTACAAATATAAAGCAATTAATTAATCAAGACACTTTATTTTCAGTT	684
QY	935	CTTATCAAGATCACGCCAAGTGTGTATTTCCCAAAAGGGCACCAAAATCGCTGGACAT	994
Db	685	CTTATCAAGATCACGCCAAGTGTGTATTTCCCAAAAGGGCACCAAAATCGCTGGACAT	744
QY	995	GTATTGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAAAGTGGAGGATTCATTTCGA	1054
Db	745	GTATTGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAAAGTGGAGGATTCATTTCGA	804
QY	1055	CCCAGATTGGCAATTTACCAAGATTTCAAGATTAAGTATTAATCTATGAAAGCTGTA	1114
Db	805	CCCAGATTGGCAATTTACCAAGATTTCAAGATTAAGTATTAATCTATGAAAGCTGTA	864
QY	1115	AGTAAACTTGGTGAAGACACAGCTACTCATTTGAAAGCCCTTGACAAATGGTCTTCCT	1174
Db	865	AGTAAACTTGGTGAAGACACAGCTACTCATTTGAAAGCCCTTGACAAATGGTCTTCCT	924
QY	1175	CCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATATCTTATAACCTGCAATCACA	1234
Db	925	CCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATATCTTATAACCTGCAATCACA	984
QY	1235	AAATAAACAATTTCTCTCATGTTTA	1261
Db	985	AAATAAACAATTTCTCTCATGTTTA	1011
RESULT	6		
AF178974			
LOCUS	AF178974	1498 bp mRNA linear	ROD 29-NOV-1999
DEFINITION	Rattus norvegicus deoxyribonuclease DLAD mRNA, complete cds.		
ACCESSION	AF178974		
VERSION	AF178974.1	GI:6470130	
KEYWORDS			
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 1498)		
AUTHORS	Tanuma, S. and Shiohara, D.		
TITLE	Cloning of a cDNA encoding a rat DNase II-like acid DNase		
JOURNAL	Biochem. Biophys. Res. Commun. 265 (2), 395-399 (1999)		
MEDLINE	20025354		
PubMed	10558878		
REFERENCE	2 (bases 1 to 1498)		
AUTHORS	Shiohara, D. and Tanuma, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-AUG-1999) Biochemistry, Science University of Tokyo, Shirojuku-ku Ichigaya Funagawara-machi 12, Tokyo 162-0826, Japan		
FEATURES	Location/Qualifiers		
source	1..1498		
	/organism="Rattus norvegicus"		
	/mol_type="mRNA"		
	/strain="Wistar"		
	/db_xref="taxon:10116"		
	159..1229		
	/codon_start=1		
	/product="deoxyribonuclease DLAD"		
	/protein_id="AA13596.1"		
	/db_xref="GI:6470131"		
CDS			

2

QY	397	AGAAAGTAGGACACACCAAAAGTTTACTGCTGTGGAAACAGAGTTCAAGGGTTCTGGGTG	456
Db	561	AGACAGTAGGACATGCGCAAAAGGTCTGCTGGTATGGAACAGAACCGAGGGTTCTGGGTG	620
QY	457	ATTCAATTCATCCCTCAGTTTCTTCCATTCGGAAGAGGCTATGATATTCACCCACA	516
Db	621	ATACACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTAGCCCACTCG	674
QY	517	GGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACGATGAGGCA	576
Db	675	GGGAGCGATATGGACAAACCGGCATCTGCATCACTTTCGGATACGCGAGTTTGAGGA	734
QY	577	ATAGATTTCTCAGCTCTTGGTCTGCAACCCCAAGGCTATAGCTGCTCATCCCGACACC	636
Db	735	ATAGATTTTCAGCTCTTGGTCTTACAACCAACATCTACAGCTGCTTCATTTCCAAGCACC	794
QY	637	TTTCCACGAGGAGCTCATTCACATGCCCGCAGCTGTGCACCGGGCCGCTCATACAGATT	696
Db	795	TTTCACTGGAACCTATYCTACATGCCCGCGATGTGCGCACTCCAGTTCTCTTAAAGATC	854
QY	697	CTGCGAGGCTCTCACCACTTCAGTCGGCGCCAGGGAACAAAATTCCTCCATTTTGCA	756
Db	855	CTGTCCGGTACCTCGCTGAACCTGCACCTCAGCCAGGCTTAACTTCGTCCATTTTGCA	914
QY	757	AAGTCGGATTCCTTCTTTGACGACATCTTTCAGCGCTGGATGGCTCAAOGGCTGAAGACA	816
Db	915	AAATCAAGTTTATATACGTATGATCTTTTACAGGATGGATAGCTCAAAAGTTGAAGACA	974
QY	817	CACTTGTTAAACAGAAACCTGGCAGCGAAAAAGACAAGAGCTTCCTTCAAAAGCTGCCCTT	876
Db	975	CAATTGTTAGCACAAACCTGGCAGAAAGAAACAAAGGCTTCTTCAAAGTTTCCCTG	1034
QY	877	CCTTACCATGCTACATATATAAAGCAATTAAATATCACGACACTCTTATTTCAAGTTCT	936
Db	1035	CCTTACCATGTCTACAACATCAAGTCATTTGGGTAACTTCCAAGTCTTACTTCAGTTCT	1094
QY	937	TATCAAGATCAACGCAAGTGTGTATTTCCCAAAAGGCAACCAAAATCCTCGACATGT	996
Db	1095	CGCCAGACCATTCCAATGGTGTGTTCCATAAAGGGCTCCGCAAAATCGCTGACCTGC	1154
QY	997	ATTGGAGACTAAATCGAGTCCACCAACAGCCTTCAGAAAGTGGAGATTCAATTGTACC	1056
Db	1155	ATTGGAGACCTAAATCGAAGCTTACCAAGCCTTAAGAGTGGAGGATTCATCTGTACA	1214
QY	1057	CAGAAATGGCAATTTTACCAAGCATTTCAAGGATAGTATTATPACTAAGAAAGCTGTAA	1116
Db	1215	AAGATCACTACATTTTACCAGGCAATTCATAAATTATATCTCGGTATGGGTTCTGTAAG	1274
QY	1117	TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACTTCACAAATGGGCTCTCTCC	1176
Db	1275	TAAACTCGGTGAAGGACCAACCCCTCTGTCTTTGAAAACTCTGGCACTGGCACTCTCG	1334
QY	1177	ATTACACC--TTCTTTATATTTTAAAGGCTGTGAAT	1211
Db	1335	CTTGGATCTGTTCTCCATAATTTTCAAGGCTTCTGAGT	1371

RESULT	8
BD091714	
LOCUS	BD091714
DEFINITION	Novel deoxyribonuclease, gene encoding thereof and use thereof.
LINEAR	linear
PATENT	PAT 27-AUG-2002

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1652)
AUTHORS	Tanuma,S. and Shikawa,D.
TITLE	Novel deoxyribonuclease, gene encoding thereof and use thereof
JOURNAL	Patent: WO 0112793-A 1 22-FEB-2001; SEICHI TANUMA,DAISUKE SHIKAWA

[illegible]

Db 1215 AAGATCACTACATTTACAGGCAATTCATAAAATATATCTCCGTTATGGGTTCTGTAAG 1274

QY 1117 TAACTTGGTGAAGGACACAGGTAATCATATCAAAACCTTGACATGGGTCTTCTTC 1176

Db 1275 TAACTCGGTGAAGGACACACCTCTGTCTTGAACACTGGCACTGGAACATCTCGC 1334

QY 1177 ATTACACC--TTCTTTATATTTTAAAGCCCTGTGAAT 1211

Db 1335 CTGGATCTGTCTCCATAATTTCAAGGCTCTGAGT 1371

RESULT 10

AR201143

LOCUS 1224 bp DNA linear PAT 20-APR-2002

DEFINITION Sequence 1 from patent US 6358723.

ACCESSION AR201143

VERSION AR201143.1

KEYWORDS GI:20252031

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1224)

AUTHORS Eastman,A.Richard, and Krieser,R.Joe.

TITLE Deoxyribonuclease II.beta. proteins and cDNAs

JOURNAL Patent: US 6358723-A 1 19-MAR-2002;

FEATURES

Location/Qualifiers

1..1224

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 53.1%; Score 672.8; DB 6; Length 1224;

Best Local Similarity 75.0%; Pred.No.3.8e-177;

Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 37 AACAGAAATGATGCAAGACTGCTAAGAACATCTTGTCTTGTCTTCTTCCCTG 96

Db 31 AGATAGAAATGACAGCAAGGCTCTAAGACAGATTCCTTCTTGTCTTCTTCCCTC 90

QY 97 TTTGGGCTCTGGGCGACCAATTTTCATCAGAAATGAGAGGAGGAAAGCTGTGGAC 156

Db 91 TCTGGGCTCTGGGCGACCAAGAAATCTCATGAGAAATGAATGTGTGAAGCTGTGGAC 150

QY 157 TGGTTACTTTTATAAGTTACTTAAAGACAAACCAAGAAAGTGGAGAGACTGGTTA 216

Db 151 TGGTTTATCTTTTATAAGTTATCCCAAGGACTAGCAAGCAAGTGAAGAGCGGGCTG 210

QY 217 GAGTACCTGTACTAGACTCTCAACTAGAACTGGAGGAGTGGAGCAACTAATGAAT 276

Db 211 CAGTACCTGTACTGGACTTCCCAAGACAAACCTGGAAACAGAGCTCTTACCTGATTAAC 270

QY 277 GACACAGAGAGTGTTTTGGAGGACATTTACACAGCTATATAGCATATGCTCTTAAG 336

Db 271 AGCACCAGAGTGTCTGGGAGGACCTTACAGCTGTGTATGACACACATAATTTCCAG 330

QY 337 AGTAACACACAGCCTATCTAATATACAAATGATGGAGTCCCTTAAACCTGTGATTTACAGT 396

Db 331 AATG---ACACAGCCTATCTAATATACACAGTGTGTCTTCCCTGATCTGTGATTTACAGC 387

QY 397 AGAAGATGACACACCAAGGTTTACTGTGTGGAAACAGATTTCAAGGGTTCTGGCTG 456

Db 388 AGACAGTATGACATGCCAAAGGCTGTGTGTATGGAACAGAAACGAGGGTTCTGGCTG 447

QY 457 ATTCAATTCATCCCTCAGTTCTCTCAATTCGGAGAGGCTATGATATCCACCCACA 516

Db 448 ATACACTGTCTTCCCAAGTTTCCCCCAGTTC-----ATGGTATGAGTACCCCACTCG 501

QY 517 GGGAGACAAATGGAACAAAGTGGCATCTGCAATACTTTCAAGTACAAACAGATGAGGCA 576

Db 502 GGGAGGCGATATGACAAACCGGATCTGCACTCTTCGGATACAGCCAGTGTGAGGAA 561

QY 577 ATAGATTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGTCTCCATCCAGCCACC 636

Db 562 ATAGATTTTTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCATTTCCAGACCC 621

QY 637 TTTCACAGAGGCTCAATTCATGCCCCAGCTGTGCACACAGGCGGCGAGCTCATTCAGAGATT 696

Db 622 TTTCACTGGAAACTTATCTACATGCCCCGAGTGTGTGCAACTCCAGTCTCTTAAAGATC 681

QY 697 CTTGGCAGGCTCTCTCACCACACTTCAGTCGGCCCGGCAAGAAATTCCTCCATTTTGA 756

Db 682 CTTGTCGGTACCTCGCTGAAGTGCATCAGCCAGGCTCTAAACTTCGCTCCATTTTGA 741

QY 757 AAGTCGGATTCTTTTCTTGACGACATCTTTTGACGCTTGGAGCTCAACCGCTGAAGACA 816

Db 742 AATCAAGTTTTTATCTATGATGATCACTTTTACAGATGGATAGCTCAAAAGTTGAAGACA 801

QY 817 CACTTGTTAACAAACCTTGGCGGCAAGAAAGCAAGAGCTTCTTCAAACTGTCCCTT 876

Db 802 CATTGTGTAGCACAACCTTGGCAGAAAAAGAAACAAGAGCTTCTTCAAACTGTCCCTG 861

QY 877 CTTTACCATGTCTACAAATATATAAAGCAATTAATATACAGACACTCTTATTTCAAGTTCT 936

Db 862 CTTTACCATGTCTACAAATATATAAAGCAATTAATATACAGACACTCTTATTTCAAGTTCT 921

QY 937 TATCAAGATCACGCCAAGTGGTGTATTTCCAAAAGGCGACCAAAATCGCTGGACATGT 996

Db 922 CGCCAAGACCATTTCCAAATGGTGTGTTCATTAAGGCTCCGCAAAATCGCTGGACCTGC 981

QY 997 ATTGGAGACTTAATTCGGAGTCCACCAAGCTTTCAGAAAGTGGAGGATTCATTTGTACC 1056

Db 982 ATTGGAGACTTAATTCGAAGCTTACCAAGCTTTCAGAAAGTGGAGGATTCATTTGTACA 1041

QY 1057 CAGAAATGGCAATTTTACCAAGCTTTCAGAAAGTGGAGGATTCATTTGTACAAGCTGTAAG 1116

Db 1042 AAGAATCACTACTATTTTACCAGGCAATTTCAATAATATATATCTCGGTTATGGTCTCTGAAG 1101

QY 1117 TAAACTTGTGTGAAGGACACAGTACTATCATTTGAAAACCTTGACAACTGGTCTTCTTCC 1176

Db 1102 TAAACTGTGTGAAGGACACACCTCTGTCTTGAACACTGGCACTGGACATCTCGC 1161

QY 1177 ATTACACTTCTTTTATATTTTA 1198

Db 1162 CTTGGATCTGTCTCCATAATA 1183

RESULT 11

AF334602

LOCUS 19164 bp DNA linear PRI 22-NOV-2001

DEFINITION Homo sapiens endonuclease DLAD (DLAD) gene, complete cds.

ACCESSION AF334602

VERSION AF334602.1

KEYWORDS GI:17046272

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 19164)

Shiokawa,D. and Tanuma,S.I.

Isolation and characterization of the DLAD/Dlad genes, which lie head-to-head with the genes for urate oxidase

Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)

21556924

PUBMED 11700027

REFERENCE 2 (bases 1 to 19164)

AUTHORS Shiokawa,D. and Tanuma,S.-I.

TITLE Direct Submission

JOURNAL Submitted (08-JAN-2001) Biochemistry, Science Univ. of Tokyo, Ichigaya, Shinjuku-ku, Tokyo 162-0826, Japan

FEATURES

Location/Qualifiers

1..19164

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1567..>17837

/gene="DLAD"

mrna join(<1567..1691,4903..5080,11439..11520,13835..13996,15315..15512,17497..>.17837)
/gene="DLAD"
/product="endonuclease DLAD"
CDS join(1567..1691,4903..5080,11439..11520,13835..13996,15315..15512,17497..17837)
/gene="DLAD"
/codon_start=1
/product="endonuclease DLAD"
/protein_id="AAL34449.1"
/db_xref="GI:17046273"
translation="MGKQWMLRTSFALLFLGLFVGLGAATFISCRNEEGKAVDMFT
FYKLPRQKESGETGLVLYLDSTTRSWRKESOLMDTKSVLGRLOQLYEAASKS
NNFAYLLYNDGVKPNYSRKYKHTGLLLNNEVQGVLIHSIPQPPPIPEEGYDVP
TGRNOSGICITFKNQVEAIDSLVNPVNSCSIPATFHOELIHPOLCTRASS
SETPGLLTLOQAOKFLHFAKSDFLDDIFAAWMAQRKLHLLTETWORQKSLP
SNGSLFYVYNIKAIKLSRHSYFSYQDHAKWCISQKTKNWTICIGDLNRSPOHQR
SGGFICTQWQIYQAFQGLVLYYESCK"

ORIGIN
Query Match 38.5%; Score 487.8; DB 9; Length 19164;
Best Local Similarity 99.6%; Pred. No. 3.2e-125;
Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 778 GACATCTTTTCAGCTGGATGGCTCAACGGCTGAAGACACACTTTGTTAACAGAAACCTGG 837
Db |||||||
QY 17496 GACATCTTTTCAGCTGGATGGCTCAACGGCTGAAGACACACTTTGTTAACAGAAACCTGG 17555
Db |||||||
QY 838 CAGCGAAAAGACAGAGCTTCTTCAAACTGCTCCCTTCTTCAACAGTCTCAATATA 897
Db |||||||
QY 17556 CAGCGAAAAGACAGAGCTTCTTCAAACTGCTCCCTTCTTCAACAGTCTCAATATA 17615
Db |||||||
QY 898 AAAGCAATTAAATATCAAGCACTCTTATTCAGTCTTATCAAGATCAAGCAAGTGG 957
Db |||||||
QY 17616 AAAGCAATTAAATATCAAGCACTCTTATTCAGTCTTATCAAGATCAAGCAAGTGG 17675
Db |||||||
QY 958 TGATTTTCCAAAAGGCGACCAAAATCGCTGGACATGATATGAGACCTTAAATCGGAGT 1017
Db |||||||
QY 17676 TGATTTTCCAAAAGGCGACCAAAATCGCTGGACATGATATGAGACCTTAAATCGGAGT 17735
Db |||||||
QY 1018 CCACACCAAGCCTTCAGAACTGGAGGATTCATTTGTACCCAGATTTGGCAATTTACCA 1077
Db |||||||
QY 17736 CCACACCAAGCCTTCAGAACTGGAGGATTCATTTGTACCCAGATTTGGCAATTTACCA 17795
Db |||||||
QY 1078 GCATTTCAAGATTAGTATTAATCAATCAAGCTGTAAGTAACTTGTGTAAGAGACACA 1137
Db |||||||
QY 17796 GCATTTCAAGATTAGTATTAATCAATCAAGCTGTAAGTAACTTGTGTAAGAGACACA 17855
Db |||||||
QY 1138 GGTACTATCATTTGAAACCTTGACAACTGGGTCTTCTTCCATTAACACCTTCTTTATTTT 1197
Db |||||||
QY 17856 GGTACTATCATTTGAAACCTTGACAACTGGGTCTTCTTCCATTAACACCTTCTTTATTTT 17915
Db |||||||
QY 1198 AAAGCGCTGTGAATATTAATTAACCTGCATATCAAAAATAAACAATTTCTCTCATG 1257
Db |||||||
QY 17916 AAAGCGCTGTGAATATTAATTAACCTGCATATCAAAAATAAACAATTTCTCTCATG 17975
Db |||||||
QY 1258 TTTACCATTTA 1268
Db |||||||
QY 17976 TTTACCATTTA 17986
Db |||||||

RESULT 12
AC027606
LOCUS Homo sapiens chromosome 1 clone RP11-30M11, WORKING DRAFT SEQUENCE,
DEFINITION 18 unordered pieces.
ACCESSION AC027606
VERSION AC027606.4 GI:9211324
KEYWORDS HTG; HRGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 164529)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 164529)
Waterston,R.H.
Direct Submission
Submitted (30-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 15, 2000 this sequence version replaced gi:8844157.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0030M11
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 156994 bases at least Q40
Consensus quality: 159185 bases at least Q30
Consensus quality: 160492 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 162829; sum-of-contigs
Quality coverage: 4.41 in Q20 bases; agarose-fp
Quality coverage: 4.23 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1304: contig of 1304 bp in length
* 1305 1404: gap of unknown length
* 1405 2815: contig of 1411 bp in length
* 2816 2915: gap of unknown length
* 2916 6065: contig of 3150 bp in length
* 6066 9367: contig of 3202 bp in length
* 9368 9467: gap of unknown length
* 9468 13313: contig of 3846 bp in length
* 13314 13413: gap of unknown length
* 13414 16861: contig of 3448 bp in length
* 16862 16961: gap of unknown length
* 16962 20927: contig of 3966 bp in length
* 20928 21027: gap of unknown length
* 21028 25671: contig of 4544 bp in length
* 25672 31485: contig of 5814 bp in length
* 31486 31585: gap of unknown length
* 31586 37141: contig of 5556 bp in length
* 37142 37241: gap of unknown length
* 37242 44119: contig of 6878 bp in length
* 44120 44219: gap of unknown length
* 44220 51993: contig of 7774 bp in length
* 51994 52093: gap of unknown length
* 52094 61678: contig of 9585 bp in length
* 61679 73742: contig of 11964 bp in length
* 73743 73842: gap of unknown length
* 73843 90362: contig of 16520 bp in length
* 90363 90462: gap of unknown length
* 90463 110115: contig of 19653 bp in length
* 110116 110215: gap of unknown length
* 110216 129841: contig of 19626 bp in length


```

* 129842 129941: gap of unknown length
* 129942 164529: contig of 34588 bp in length.
FEATURES
    source
        1. 164529
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-30M11"
            1. 1304
            /note="assembly_name:Contig11"
            1405..2815
            /note="assembly_name:Contig12"
            clone_end:SP6
            vector_side:right
            2916..6065
            /note="assembly_name:Contig13"
            6166..9367
            /note="assembly_name:Contig14"
            9468..13313
            /note="assembly_name:Contig15"
            13414..16861
            /note="assembly_name:Contig16"
            16962..20927
            /note="assembly_name:Contig17"
            21028..25571
            /note="assembly_name:Contig18"
            25672..31485
            /note="assembly_name:Contig19"
            31586..37141
            /note="assembly_name:Contig20"
            37242..44119
            /note="assembly_name:Contig21"
            44220..51993
            /note="assembly_name:Contig22"
            52094..61678
            /note="assembly_name:Contig23"
            61779..73742
            /note="assembly_name:Contig24"
            73843..90362
            /note="assembly_name:Contig25"
            90463..110115
            /note="assembly_name:Contig26"
            110216..129841
            /note="assembly_name:Contig27"
            129942..164529
            /note="assembly_name:Contig28"

ORIGIN
Query Match      38.5%; Score 487.8; DB 2; Length 164529;
Best Local Similarity 99.6%; Pred. No. 3.8e-125;
Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 778 GACATCTTGCAGCTGGATGCTCAACGGCTGAGACACATCTGTTAAAGAAACCTGG 837
DB 143110 GACATCTTGCAGCTGGATGCTCAACGGCTGAGACACATCTGTTAAAGAAACCTGG 143169

QY 838 CAGCGAAAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTACAAATATA 897
DB 1431170 CAGCGAAAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTACAAATATA 143229

QY 898 AAAGCAATTAATATCATCGACATCTTATTTCAGTCTCTTATCAAGATCACGCCAAGTGG 957
DB 143230 AAAGCAATTAATATCATCGACATCTTATTTCAGTCTCTTATCAAGATCACGCCAAGTGG 143289

QY 958 TGTATTTCCCAAGGGCACCAAAATCGCTGGACATGATTTGGAGACTAAATCGGAGT 1017
DB 143230 TGTATTTCCCAAGGGCACCAAAATCGCTGGACATGATTTGGAGACTAAATCGGAGT 143349

QY 1018 CCACACCAAGCTTCAGAAAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 1077
DB 143350 CCACACCAAGCTTCAGAAAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 143409

1078 GCATTTCAGGATAGTATTATATCTATGAAAGCTGTAAAGTAAACCTGTGTGAAGACACA 1137
143410 GCATTTCAGGATAGTATTATATCTATGAAAGCTGTAAAGTAAACCTGTGTGAAGACACA 143469

1138 GGTACTATCATTTGAAAACCTTGCAATGGGTCTTCTCCATTACACCTCTCTTATATTTT 1197
143470 GGTACTATCATTTGAAAACCTTGCAATGGGTCTTCTCCATTACACCTCTCTTATATTTT 143529

1198 AAGGCTCTGTAATATATCTTATTAACCTGCAATATACAAAATAAATATTTCTCTCATG 1257
143530 AAGGCTCTGTAATATATCTTATTAACCTGCAATATACAAAATAAATATTTCTCTCATG 143589

1258 TTTACCAATTTA 1268
143590 TTTACCAATTTA 143600

RESULT 13
AL359273
LOCUS
DEFINITION
    Human DNA sequence from clone RP11-376N17 on chromosome 1, complete
    sequence.
ACCESSION
    AL359273
VERSION
    AL359273.11 GI:12191184
KEYWORDS
    HTG.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 205034)
AUTHORS
    Wallis,J.
TITLE
    Direct Submission
JOURNAL
    Submitted (10-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
    CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
    requests: clonerequest@sanger.ac.uk
COMMENT
    On Jan 13, 2001 this sequence version replaced gi:12043432.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission
    corresponding to the overlapping clone, as we submit sequences with
    only a small overlap as described above.
    This sequence has been finished according to sequence map criteria
    as follows. An attempt is made to resolve all sequencing problems,
    such as compressions and repeats, but not necessarily within known
    annotated repeat sequence elements. Where the sequence is
    ambiguous, there is an annotation using the 'unsure' feature key.
    The following abbreviations are used to associate primary accession
    numbers given in the feature table with their source databases:
    Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
    on the WORMPEP database can be found at
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep
    This sequence
    was generated from part of bacterial clone contigs of human
    chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
    Group. Further information can be found at
    http://www.sanger.ac.uk/HGP/Chri
    RP11-376N17 is from the library RPCI-11.2 constructed by the group
    of Pieter de Jong. For further details see
    http://www.chori.org/bacpac/home.htm
    VECT01: pBACe3.6
    This sequence is the entire insert of clone RP11-376N17 The true
    right end of clone RP11-82H13 is at 76245 in this sequence.
    location/Qualifiers
        1. 205034
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-376N17"
            /clone_lib="RPCI-11.2"
            512..622
            /note="L2 repeat: matches 2588. 2704 of consensus"
            764..845
            repeat_region
            repeat_region

```

```

repeat_region /note="41 copies 2 mer ct 68% conserved"
repeat_region 2021. .2169
repeat_region /note="MIR repeat: matches 58. .210 of consensus"
repeat_region 2802. .2886
repeat_region /note="MIR repeat: matches 102. .192 of consensus"
repeat_region 3607. .4457
repeat_region /note="L1MB3 repeat: matches 5284. .6181 of consensus"
repeat_region 4671. .4952
repeat_region /note="AluSx repeat: matches 1. .301 of consensus"
repeat_region 6766. .7048
repeat_region /note="AluSx repeat: matches 1. .307 of consensus"
repeat_region 7636. .7693
repeat_region /note="29 copies 2 mer at 87% conserved"
repeat_region 8235. .8452
repeat_region /note="MER30 repeat: matches 1. .230 of consensus"
repeat_region 8483. .8592
repeat_region /note="MIR repeat: matches 67. .166 of consensus"
repeat_region 8578. .8717
repeat_region /note="MIR repeat: matches 118. .257 of consensus"
repeat_region 8887. .9474
repeat_region /note="L2 repeat: matches 1728. .2356 of consensus"
repeat_region 9508. .10010
repeat_region /note="MLTIF repeat: matches 1. .541 of consensus"
repeat_region 10247. .10349
repeat_region /note="MLTID repeat: matches 23. .120 of consensus"
repeat_region 10351. .10453
repeat_region /note="L1MB6 repeat: matches 6074. .6176 of consensus"
repeat_region 10457. .10693
repeat_region /note="MER7A repeat: matches 1. .223 of consensus"
repeat_region 10694. .10817
repeat_region /note="L1MB7 repeat: matches 6159. .6283 of consensus"
repeat_region 10838. .10883
repeat_region /note="23 copies 2 mer ta 76% conserved"
repeat_region 10908. .11211
repeat_region /note="MLTID repeat: matches 192. .505 of consensus"
repeat_region 11215. .11380
repeat_region /note="L2 repeat: matches 2555. .2749 of consensus"
repeat_region 11742. .11813
repeat_region /note="MIR repeat: matches 65. .145 of consensus"
repeat_region 11838. .12325
repeat_region /note="L2 repeat: matches 2245. .2746 of consensus"
repeat_region 12326. .12658
repeat_region /note="MER33 repeat: matches 5. .324 of consensus"
repeat_region 12659. .12770
repeat_region /note="L2 repeat: matches 2131. .2245 of consensus"
repeat_region 12771. .13068
repeat_region /note="AluY repeat: matches 1. .294 of consensus"
repeat_region 13069. .13324
repeat_region /note="L2 repeat: matches 1895. .2131 of consensus"
repeat_region 13410. .13524
repeat_region /note="L2 repeat: matches 2269. .2384 of consensus"
repeat_region 13668. .14270
repeat_region /note="MER67B repeat: matches 1. .661 of consensus"
repeat_region 14405. .14569
repeat_region /note="MER69A repeat: matches 1. .163 of consensus"
repeat_region 14842. .16800
repeat_region /note="L1MC1 repeat: matches 4339. .6327 of consensus"
repeat_region 16828. .16920
repeat_region /note="MSTD repeat: matches 41. .132 of consensus"
repeat_region 17020. .17238
repeat_region /note="MSTD repeat: matches 199. .426 of consensus"
repeat_region 17239. .17656
repeat_region /note="L1MC1 repeat: matches 3954. .4343 of consensus"
repeat_region 17650. .18127
repeat_region /note="L1MCa repeat: matches 329. .831 of consensus"
repeat_region 18128. .18405
repeat_region /note="AluJb repeat: matches 3. .278 of consensus"
repeat_region 18406. .18515
repeat_region /note="L1MCa repeat: matches 224. .329 of consensus"
repeat_region 18516. .18600
repeat_region /note="MER5A repeat: matches 67. .162 of consensus"
repeat_region 20386. .20677
repeat_region /note="AluSg repeat: matches 1. .300 of consensus"

```

```

repeat_region 21518. .21560
repeat_region /note="L2 repeat: matches 2624. .2667 of consensus"
repeat_region 21587. .21746
repeat_region /note="MIR repeat: matches 67. .231 of consensus"
repeat_region 21847. .22101
repeat_region /note="L2 repeat: matches 2244. .2489 of consensus"
repeat_region 22265. .24015
repeat_region /note="L1MC1 repeat: matches 2197. .3954 of consensus"
repeat_region 24177. .24480
repeat_region /note="AluSg repeat: matches 6. .301 of consensus"
repeat_region 24979. .25248
repeat_region /note="MLTII repeat: matches 63. .321 of consensus"
repeat_region 25329. .25560
repeat_region /note="HALI repeat: matches 557. .792 of consensus"
repeat_region 25590. .26037
repeat_region /note="L1MEC repeat: matches 1207. .1648 of consensus"
repeat_region 27307. .27752
repeat_region /note="MLTIC repeat: matches 1. .466 of consensus"
repeat_region 28933. .29100
repeat_region /note="HALI repeat: matches 585. .772 of consensus"
repeat_region 29101. .30182
repeat_region /note="MER11C repeat: matches 1. .1071 of consensus"
repeat_region 30183. .30284
repeat_region /note="HALI repeat: matches 772. .868 of consensus"
repeat_region 30670. .31206
repeat_region /note="MLTIF repeat: matches 1. .538 of consensus"
repeat_region 31207. .31928
repeat_region /note="MER67B repeat: matches 2. .661 of consensus"
repeat_region /note="complement (31605. .32107)"
repeat_region /note="match: STS: Em:HSP16C04"
repeat_region 31931. .32036
repeat_region /note="53 copies 2 mer aa 63% conserved"
repeat_region 32476. .32893
repeat_region /note="MSTA repeat: matches 2. .425 of consensus"
repeat_region 33144. .33452
repeat_region /note="AluSx repeat: matches 1. .310 of consensus"
repeat_region 34297. .34983
repeat_region /note="L2 repeat: matches 1992. .2750 of consensus"
repeat_region 35174. .35467
repeat_region /note="AluSg repeat: matches 3. .297 of consensus"
repeat_region 35900. .36248
repeat_region /note="THRIC repeat: matches 3. .371 of consensus"
repeat_region 36615. .36696
repeat_region /note="L2 repeat: matches 1711. .1779 of consensus"
repeat_region 36697. .39050
repeat_region /note="TIGER1 repeat: matches 3. .2418 of consensus"
repeat_region 39051. .39300
repeat_region /note="L2 repeat: matches 1779. .2094 of consensus"
repeat_region 39347. .39578
repeat_region /note="MIR repeat: matches 23. .262 of consensus"
repeat_region 39897. .39988
repeat_region /note="AluSg/x repeat: matches 218. .309 of consensus"
repeat_region 40510. .41307
repeat_region /note="CpG island"
repeat_region /evidence=not_experimental
repeat_region 41742. .41781
repeat_region /note="20 copies 2 mer ta 87% conserved"
repeat_region 43592. .44272
repeat_region /note="match: STS: Em:G51282"
repeat_region 44412. .44473
repeat_region /note="31 copies 2 mer tt 67% conserved"
repeat_region 44611. .44660
repeat_region /note="25 copies 2 mer ca 86% conserved"
repeat_region 46419. .46584
repeat_region /note="MER3 repeat: matches 1. .164 of consensus"
repeat_region 46591. .48600
repeat_region /note="L1ME1 repeat: matches 4097. .6163 of consensus"
repeat_region 47717. .47748
repeat_region /note="Sequence from overlapping clone bA82H13 (AL450063). Assembly confirmed by restriction digest data."
repeat_region 48601. .48916
repeat_region /note="AluSg repeat: matches 1. .287 of consensus"
repeat_region 48922. .48967

```

```

Query Match      38.5%; Score 487.8; DB 9; Length 205034;
Best Local Similarity 99.6%; Pred. No. 3.8e-125;
Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 778 GACATCTTTGAGCTGAGCTGCTCAACGGCTGAGACACACTTGTGTTAACGAAACCTGG 837
Db 156611 GACATCTTTGAGCTGAGCTGCTCAACGGCTGAGACACACTTGTGTTAACGAAACCTGG 156670

QY 838 CAGCGAAAAAGACAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATGTTACAATATA 897
Db 156671 CAGCGAAAAAGACAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATGTTACAATATA 156730

QY 898 AAGCAATTAATATCAAGACACTCTTATTTCAAGTTCCTTATCAAGATCAAGCAAGTGG 957
Db 156731 AAGCAATTAATATCAAGACACTCTTATTTCAAGTTCCTTATCAAGATCAAGCAAGTGG 156790

QY 958 TGATTTTCCCAAGGGCACCAGAAATCGCTGGACATGTTAGAGACCTTAAATCGGAGT 1017
Db 156791 TGATTTTCCCAAGGGCACCAGAAATCGCTGGACATGTTAGAGACCTTAAATCGGAGT 156850

QY 1018 CCACACCAAGCTTCAAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAAATTTACCAA 1077
Db 156851 CCACACCAAGCTTCAAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAAATTTACCAA 156910

QY 1078 GCATTTTCAAGGATAGTATTAATCACTATGAAGCTGTAAGTAAACTTGGTGAAGGACACA 1137
Db 156911 GCATTTTCAAGGATAGTATTAATCACTATGAAGCTGTAAGTAAACTTGGTGAAGGACACA 156970

QY 1138 GGTACTATCATTCAGAAACCTTGCAATGGTCTTCTTCCATTCACCTTCTTTATATTTT 1197
Db 156971 GGTACTATCATTCAGAAACCTTGCAATGGTCTTCTTCCATTCACCTTCTTTATATTTT 157030

QY 1198 AAGGGCTGTGAATATCTATTAACCTGCATATCAGAAATTAACATATTTCTTCATG 1257
Db 157031 AAGGGCTGTGAATATCTATTAACCTGCATATCAGAAATTAACATATTTCTTCATG 157090

QY 1258 TTTACCAATTA 1268
Db 157091 TTTACCAATTA 157101

RESULT 14
AF334603S6 2337 bp DNA linear ROD 22-NOV-2001
LOCUS Mus musculus deoxyribonuclease DLAD (Dlad) gene, exon 6 and complete cds.
ACCESSION AF334608
VERSION AF334608.1 GI:17046280
SEGMENT 6 of 6
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2337)
Isolation and characterization of the DLAD/Dlad genes, which lie head-to-head with the genes for urate oxidase
Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)
21556924
PUBMED 11700027
REFERENCE 2 (bases 1 to 2337)
AUTHORS Shiohawa,D. and Tanuma,S.-I.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2001) Department of Biochemistry, Science
JOURNAL University of Tokyo, Ichigaya, Shinjuku-ku, Tokyo 162-0826, Japan
FEATURES
Location/Qualifiers
1..2337
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6"
/db_xref="taxon:10090"

```

```

order(AF334603.1:1476..1825,AF334604.1:1..1374,
AF334605.1:1..1089,AF334606.1:1..719,AF334607.1:1..642,
1..1182)
/gene="Dlad"
join(AF334603.1:1476..1588,AF334604.1:665..842,
AF334605.1:630..708,AF334606.1:484..639,
AF334607.1:102..299,486..1182)
/gene="Dlad"
/product="deoxyribonuclease DLAD"
join(AF334603.1:1476..1588,AF334604.1:665..842,
AF334605.1:630..708,AF334606.1:484..639,
AF334607.1:102..299,486..826)
/gene="Dlad"
/codon_start=1
/product="deoxyribonuclease DLAD"
/protein_id="AAL34450.1"
/db_xref="GI:17046281"
/translation="MTAKPLRTVLSLFFALSGVLGPEISCRNEYGEAVDFYFKL
PKTSKASEAGLQYIYLDSTROTWNKSLYLINSTRALGRTIQLHYDTHSINDTAY
LIYNDGVPGSVNSROYHAKGLLVNRTQGFWLIIHSPKFPVPHGVYESTSGRRYQG
TGCTTFGYSQFEEDFOLLVLPNIYSCFIPSTFWKLTYPMRMCANSLSKIPVRY
LAELHQAQGLNFVHFAKSFYTDIDFTGMLAOKLTHLLAOTWQKKQLPNSCSLPY
KHNYSIGVTSKSYFSSRDHSKWCVSIGSANRWTCTGDLNRSLSHQLRGGGFCIT
KNHYIYQAFHKLILRYGCK"
486..1182
/gene="Dlad"
/number=6

Query Match      20.0%; Score 254; DB 10; Length 2337;
Best Local Similarity 75.5%; Pred. No. 9.4e-60;
Matches 329; Conservative 0; Mismatches 105; Indels 2; Gaps 1;

QY 778 GACATCTTTGAGCTGAGTGGCTCAACGGCTGAGACACACTTGTGTTAACGAAACCTGG 837
Db 485 GACATCTTTACAGGATGAGTAGCTCAAAAGTTGAGACACATTTGTAGCACAAACCTGG 544

QY 838 CAGCGAAAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGTTACAATATA 897
Db 545 CAGAAAAAGAAAAAGAGCTTCTTCAAACTGTTCTCGCTTACCATGTTACAATATA 604

QY 898 AAGCAATTAATATCAAGACACTCTTATTTCAAGTTCCTTATCAAGATCAAGCAAGTGG 957
Db 605 AAGTCAATTTGGGGTAACCTTCCAAAGCTTCTTACCTTCTCGCCAGACCATTCCTAATGG 664

QY 958 TGATTTTCCCAAGGGCACCAGAAATCGCTGACATGTTATGGAGACCTTAATCGGAGT 1017
Db 665 TGTTGTTTCCATAAAGGGCTCCGCAATCGCTGACCTGCAITGGAGACCTTAATCGAAGC 724

QY 1018 CCACACCAAGCTTCAAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAAATTTACCAA 1077
Db 725 CTACACCAAGCTTAAAGAGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACGAG 784

QY 1078 GCATTTTCAAGGATAGTATTAATCACTATCAAGAGCTGTAAGTAAACTTGGTGAAGGACACA 1137
Db 785 GCATTTTCAATTAATATATATCTCCGTTATGGGTCTGTAAAGTAACTCGGTGAAAGGCCACA 844

QY 1138 GGTACTATCATTCAGAAACCTTGCAATGGTCTTCTTCCATTTACACC--TTCTTTATATT 1195
Db 845 CCTCTGTCTTGAAGAACACTGGCACTGGACATCTCGCCTTGGATCTGTCTTCCATTAAT 904

QY 1196 TTAAGGCTGTGAAT 1211
Db 905 TTCAAGGCTTCTGAGT 920

RESULT 15
AC127571 197457 bp DNA linear HTG 17-OCT-2002
LOCUS Mus musculus chromosome UNK clone RP24-227D4, WORKING DRAFT
DEFINITION SEQUENCE, 4 unoriented pieces.
ACCESSION AC127571
VERSION AC127571.2 GI:24080752

```

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)
ORGANISM
Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 197457)
McPherson, J.D. and Waterston, R.H.
TITLE
The sequence of Mus musculus clone
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 197457)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
3 (bases 1 to 197457)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (17-OCT-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT
On Oct 17, 2002 this sequence version replaced gi:21887002.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M BB0227D04
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 195668 bases at least Q40
Consensus quality: 196117 bases at least Q30
Consensus quality: 196344 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31865: contig of 31865 bp in length
* 31866 31865: gap of unknown length
* 31966 70958: contig of 38993 bp in length
* 70959 71058: gap of unknown length
* 71059 136518: contig of 65460 bp in length
* 136519 136618: gap of unknown length
* 136619 197457: contig of 60839 bp in length.
FEATURES
source
1..197457
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-227D4"
misc_feature
1..31865
/note="assembly_name:Contig14"
misc_feature
31966..70958
/note="assembly_name:Contig15"
misc_feature
71059..136518
/note="assembly_name:Contig16"
misc_feature
136619..197457
/note="assembly_name:Contig17"
ORIGIN
Query Match 20.0%; Score 254; DB 2; Length 197457;
Best Local Similarity 75.5%; Pred. No. 1.3e-59;

Matches 329; Conservative 0; Mismatches 105; Indels 2; Gaps 1;
QY 778 GACATCTTTGACGCTGGATGGCTCAACGGCTGAAGACACACCTGTGTAAAGAAACCTGG 837
Db 68107 GACATCTTTTACAGGATGGATAGCTCAAAGTTGAAGACACACATTTGTAGCACAAACCTGG 68166
QY 838 CAGCGAAAAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTCTACATATA 897
Db 68167 CAGAAAAAGAAACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTCTACATATA 68226
QY 898 AAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 957
Db 68227 AAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCAGTTCTCGCCCAAGACCAATTCGAATGG 68286
QY 958 TGTATTTTCCCAAAAGGGCACCAAAATTCGCTGACATGTATTTGGAGACCTTAATTCGAATGG 1017
Db 68287 TGTATTTTCCCAAAAGGGCACCAAAATTCGCTGACATGTATTTGGAGACCTTAATTCGAATGG 68346
QY 1018 CCAACCAAGCTTCAGAGTGGAGATTCATTTGTATCCAGATTTGGCAATTTACCAA 1077
Db 68347 CTACACCAAGCTTAAAGAGTGGAGATTCATCTGTACAAAGATTCATCTACATTTACCA 68406
QY 1078 GCATTTCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1137
Db 68407 GCATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 68466
QY 1138 GGTACTATCATTTGAAAACCTTTGACATGGGTCTTTTCCATTAACCC--TTCTTTTATTT 1195
Db 68467 CCCTCTGCTCTTGAACACACTGGCAGTGAACATCTCGCTTGGATCTGTTTCTCCATAAT 68526
QY 1196 TTAAGGCTGTGAAT 1211
Db 68527 TTCAAGGCTTCTGAGT 68542
Search completed: October 14, 2004, 03:42:41
Job time : 4901.47 secs

Result	Query			DB	ID	Description
	No.	Score	Match			
1	1268	100.0	1268	5	AAD21289	Human deo
2	1260.4	99.4	1690	5	ABV29663	Human pro
3	1260.4	99.4	1690	5	ABV29678	Human pro
4	1140	89.9	1140	6	ABK92136	Prostate
5	1079.6	85.1	1086	4	AAF60716	Human DNA
6	673.8	53.1	1652	4	AP60715	Murine DNA
7	672.8	53.1	1224	5	AAD21288	Mouse deo
8	310.8	24.5	480	8	ACH19937	Human adu
9	155	12.2	266145	9	ADE87477	Fowlbox v
10	84	6.6	1915	2	AAV29137	Homo sapi
11	82.4	6.5	1575	2	AAV02903	Human DNA
12	71	5.6	927	2	AAV29138	Homo sapi
13	60	4.7	60	6	ABN04074	Bos tauru
14	57	4.5	65	6	ABN04074	Human spl
15	54.4	4.3	544	7	ACD98178	Rat splic
16	50.6	4.0	478	8	ACH13974	Human col
17	46.6	3.7	222	7	ABX55053	Human adu
18	45	3.5	2000	7	ADA71938	Bovine ES
19	43.4	3.4	2000	7	ADA71938	Rice gene
20	39	3.1	169998	6	AD36511	Rice gene
21	39	3.1	197496	6	ABN85584	Human Hgr
22	38.6	3.0	8759	7	ABZ10109	Human EGF
23	38.6	3.0	8759	7	ABZ10237	Haematopo

The invention relates to deoxyribonuclease (DNase) II beta proteins and their corresponding cDNAs. The DNase II beta may be useful to digest DNA in the mucous plugs in lungs of cystic fibrosis patients and so reduce their viscosity. The present sequence is human DNase II beta cDNA. The

961 ATTTCCCAAAGGCACCAAAATCGCTGGACATGTATTGGAGACCTAAATCGGAGTCCA 1020

determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Query Match 99.4%; Score 1260.4; DB 5; Length 1690;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

7 AAAGTGTCTGCTGGCATGAAAATAAATGAACAGAAAAATGATGCCAGACTGCTTAAGA 66
|||
7 AAAGTGTCTGCTGGCATGAAAATAAATGAACAGAAAAATGATGCCAGACTGCTTAAGA 66
|||

67 ACATTCCTTTTGCTTGTCTTCTTGGCCTCTTTGGGGTGCCTGGGGCGACCAACAATTCCA 126
|||
67 ACATTCCTTTTGCTTGTCTTCTTGGCCTCTTTGGGGTGCCTGGGGCGACCAACAATTCCA 126
|||

127 TGCAGAAATGAAGAGGGAAGCTGTGACHTGGTCTTCTTTTATATAGTTACTTAAAGA 186
|||
127 TGCAGAAATGAAGAGGGAAGCTGTGACHTGGTCTTCTTTTATATAGTTACTTAAAGA 186
|||

187 CAAAAACAAGGAAAGTGGAGAGACTGGGTGTAGAGTACCCTGTACCTTAGACTCTTACAACTAGA 246
|||
187 CAAAAACAAGGAAAGTGGAGAGACTGGGTGTAGAGTACCCTGTACCTTAGACTCTTACAACTAGA 246
|||

247 AGCTGGAGGGAAGAGTGTGACCACTAATGAATGACACCAAGAGTGTTTGGGAAGGACAAATTA 306
|||
247 AGCTGGAGGGAAGAGTGTGACCACTAATGAATGACACCAAGAGTGTTTGGGAAGGACAAATTA 306
|||

307 CAACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACAACAGCCTATCTAATATACAAT 366
|||
307 CAACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACAACAGCCTATCTAATATACAAT 366
|||

367 GATGGAGTCCCTAAACCTGTGAATTAACGTAGAAAAGTATGGAACAACAAGGTTTACTG 426
|||
367 GATGGAGTCCCTAAACCTGTGAATTAACGTAGAAAAGTATGGAACAACAAGGTTTACTG 426
|||

427 CTGTGGAACAAGAGTCAAGGGTTCTGGCTCATTCATTCCTCAATCCCTCAGTTTCTCCCAATT 486
|||
427 CTGTGGAACAAGAGTCAAGGGTTCTGGCTCATTCATTCCTCAATCCCTCAGTTTCTCCCAATT 486
|||

487 CCAGGAAGAGGCTATGATTAATCCACCAACAGGAGAGCAAAATGGAACAAAGTGCATCTGC 546
|||
487 CCAGGAAGAGGCTATGATTAATCCACCAACAGGAGAGCAAAATGGAACAAAGTGCATCTGC 546
|||

547 ATAACTTTCAAGTACAAACAGTATAGGCAATAGATTCTCAGCTCTTGGTCTGCAACCCC 606
|||
547 ATAACTTTCAAGTACAAACAGTATAGGCAATAGATTCTCAGCTCTTGGTCTGCAACCCC 606
|||

607 AACGCTCTATAGCTGCTCCATCCAGCACCTTTCACAGGAGTCAATTCATGCCCCAG 666
|||
607 AACGCTCTATAGCTGCTCCATCCAGCACCTTTCACAGGAGTCAATTCATGCCCCAG 666
|||

667 CTGTGCAACAGGCGCAGCTCATCAGAGATTCTGGGAGGCTCTCTACCACTTCACTGCTG 726
|||
667 CTGTGCAACAGGCGCAGCTCATCAGAGATTCTGGGAGGCTCTCTACCACTTCACTGCTG 726
|||

727 GCCCAGGGAACAAAATTCCTCCATTTGCAAAGTCGGAATCTTTTCTTGACGACATCTTT 786
|||
727 GCCCAGGGAACAAAATTCCTCCATTTGCAAAGTCGGAATCTTTTCTTGACGACATCTTT 786
|||

787 GCAGCTGTGATGGCTCAACGGCTGAAGAACAACCTGTTAACAGAAACCTGGCAGCGAAA 846
|||
787 GCAGCTGTGATGGCTCAACGGCTGAAGAACAACCTGTTAACAGAAACCTGGCAGCGAAA 846
|||

847 AGACAAGAGCTTCCCTTCAAACTGCTCCCTTCTTACCATGTCTACAATATAAAGCAATT 906
|||
847 AGACAAGAGCTTCCCTTCAAACTGCTCCCTTCTTACCATGTCTACAATATAAAGCAATT 906
|||

907 AAAATTATCAGCACACTTATTTTCAAGTCTTATCAAGATCATCGCAAGTGGTGTATTTCC 966
|||
907 AAAATTATCAGCACACTTATTTTCAAGTCTTATCAAGATCATCGCAAGTGGTGTATTTCC 966
|||

QY	967	CAAAAGGCGCACCAAAAATCGCTGCACATGTATTGGAGACCTTAAATCGAGTCCACACAA	1026
Db	967	CAAAAGGCGCACCAAAAATCGCTGCACATGTATTGGAGACCTTAAATCGAGTCCACACAA	1026
QY	1027	GCCTTCAGAAGTGGAGGATTCATTTTGTACCCAGAATTTGGCAAAATTTACCAAGCATTTTCAA	1086
Db	1027	GCCTTCAGAAGTGGAGGATTCATTTTGTACCCAGAATTTGGCAAAATTTACCAAGCATTTTCAA	1086
QY	1087	GGATTAGTATTATATCTATGAAAGCTGTAAAGTAAACTTTGGTGAAGGACACAGGTACTATC	1146
Db	1087	GGATTAGTATTATATCTATGAAAGCTGTAAAGTAAACTTTGGTGAAGGACACAGGTACTATC	1146
QY	1147	ATTGAAAACCTTGCACATGGCTCTTCTTCCATTACACCTTCTTTATATTTTAAAGGCGCTG	1206
Db	1147	ATTGAAAACCTTGCACATGGCTCTTCTTCCATTACACCTTCTTTATATTTTAAAGGCGCTG	1206
QY	1207	TGAATATCTATTAAACCTTGCATATCAAAAATAAAACATATTTCTCTCATGTTTACCATT	1266
Db	1207	TGAATATCTATTAAACCTTGCATATCAAAAATAAAACATATTTCTCTCATGTTTACCATT	1266
QY	1267	TA 1268	
Db	1267	TA 1268	

RESULT 3

ABV23788
ID ABV23788 standard; cDNA; 1690 BP.

XX AC ABV23788;

XX
DT 16-SEP-2002 (first entry)

XX
DE
....
Human prostate expression marker cDNA 23779.

xx
xx
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.XX
PN
WO200160860-A2

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001WO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
DR WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 4376; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC	determining whether prostate cancer has metastasized in a patient; (h)
CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX	
SQ	Sequence 1690 BP; 525 A; 369 C; 305 G; 485 T; 0 U; 6 Other;
	Query Match 99.4%; Score 1260.4; DB 5; Length 1690;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	7 AAAGTGTCTCGTCGTGGCATGAAATAAATGAAGAACAAGAAATGATGGCAAGACTGCCTAAGA 66
DB	7 AAAGTGTCTCGTCGTGGCATGAAATAAATGAACACAGAAAATGATGGCAAGACTGCCTAAGA 66
QY	67 ACATCCTTTTGCTTTGCTCTTCCTTGGGCTCTTTGGGGTGTCTGGGGGAGCAAAACAATTCA 126
DB	67 ACATCCTTTTGCTTTGCTCTTCCTTGGGCTCTTTGGGGTGTCTGGGGGAGCAACAATTCA 126
QY	127 TGCAGAAATCAAGAAGGAAAGCTGTGGACCTGGTTTATCTTTTTATAAGTTACCCTAAAGA 186
DB	127 TGCAGAAATGAGAGAGGAAAGCTGTGGACCTGGTTTACTTTTTATAAGTTACCTAAAGA 186
QY	187 CAATAACAAGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTACAACTAGA 246
DB	187 CAATAACAAGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTTACAACTAGA 246
QY	247 AGCTGGAGGAGAGTGGACACTTAATGAATCACACCAAGAGTGTTTGGGAAGGACATTA 306
DB	247 AGCTGGAGGAGAGTGGACACTTAATGAATGACACCAAGAGTGTTTGGGAGAGGACATTA 306
QY	307 CAACAGCTTATATGAAGCATATGCCCTCTAAGAGTAAACAACAAGGCTATCTAATATACAAT 366
DB	307 CAACAGCTTATATGAAGCATATGCCCTCTAAGAGTAAACAACAAGGCTATCTAATATACAAT 366
QY	367 GATGGAGTCCCATAAACCCTGTGAATTPACAGTGAAGAAGTATGAGACACACCAAAGGTTTACTG 426
DB	367 GATGGAGTCCCATAAACCTGTGAATTTACAGTGAAGAAGTATGAGACACACCAAAGGTTTACTG 426
QY	427 CTGTGGAAACAGAGTTCAGAGGTTCTGGCTGATTCATTTCCATCCCTCAGTTTCCTCCAATT 486
DB	427 CTGTGGAAACAGAGTTCAGAGGTTCTGGCTGATTCATTTCCATCCCTCAGTTTCCTCCAATT 486
QY	487 CGGGAAGAGGCTATGATTTATCCACCACAGGGAGACGAAATGGACAAAGTGGCATCTGC 546
DB	487 CGGGAAGAGGCTATGATTTATCCACCAAGGGAGACGAAATGGACAAAGTGGCATCTGC 546
QY	547 ATAACTTTTCAAGTACAAACGAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGGCAACCCC 606
DB	547 ATAACTTTTCAAGTACAAACGAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGGCAACCCC 606
QY	607 AACGCTATATAGTGTCTCCATCCAGGCACCTTTACACAGGAGCTCATTCACATGCCCCAG 666
DB	607 AACGCTATATAGTGTCTCCATCCAGGCACCTTTACACAGGAGCTCATTCACATGCCCCAG 666
QY	667 CTGTGCACACAGGGCCAGCTCATCAGAGATTCCTTGGCAGGCTCCTCACCACTTCAGTCG 726
DB	667 CTGTGCACACAGGGCCAGCTCATCAGAGATTCCTTGGCAGGCTCCTCACCACTTCAGTCG 726
QY	727 GCCACGGGACAAAAATTCCTCCATTTTGCATTTGCAAGTGGATCTTTTCTTGACCACTCTTT 786
DB	727 GCCACGGGACAAAAATTCCTCCATTTTGCATTTGCAAGTGGATCTTTTCTTGACCACTCTTT 786
QY	787 GCAGCTGATGGCTCAACGGGTGAAGACACACTTCTTAAACAGAAACCTGGCAGCGAAAA 846
DB	787 GCAGCTGATGGCTCAACGGGTGAAGACACACTTGTTLACAGAAACCTGGCAGCGAAAA 846
QY	847 AGCAAGAGGTTCCCTTCAAACCTGCTCCCTTCCCTTACCATGTCTACAAATATAAAGCAATT 906
DB	847 AGCAAGAGGTTCCCTTCAAACCTGCTCCCTTCCCTTACCATGTCTACAAATATAAAGCAATT 906
QY	907 AAAATTATCAGCACTCTTATTTCAAGTTCCTTATCAAGATCAACGCGAGTGTGTTATTTCC 956

XX The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridize to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
 CC sequences
 CC
 XX Sequence 1140 BP; 357 A; 258 C; 245 G; 280 T; 0 U; 0 Other;
 SQ
 Query Match 89.9%; Score 1140; DB 6; Length 1140;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGGGAAAGTGTCTCTGCTGGCATGAATAAATGAACAGAAATGATGGCAAGACTG 60
 1 ATGGGAAAGTGTCTCTGCTGGCATGAATAAATGAACAGAAATGATGGCAAGACTG 60
 61 CTAAGAACATCTTCTGCT 120
 61 CTAAGAACATCTTCTGCT 120
 121 ATTTTCATGAGAAATGAAGAGGAAAGCTGTGGACTGTGGTCTCTCTCTCTCTCTCTCT 180
 121 ATTTTCATGAGAAATGAAGAGGAAAGCTGTGGACTGTGGTCTCTCTCTCTCTCTCTCT 180
 181 AAAAGACAAAACAGGAAAGTGGAGAGTGGGTTAGAGTACTGTACCTAGACTCTACA 240
 181 AAAAGACAAAACAGGAAAGTGGAGAGTGGGTTAGAGTACTGTACCTAGACTCTACA 240
 241 ACTAGAGCTGGAGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAG 300
 241 ACTAGAGCTGGAGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAG 300
 301 ACATTACACACATATGAAGCATATGCTCTTAAGAGTAAACACACAGCCTATCTAATA 360
 301 ACATTACACACATATGAAGCATATGCTCTTAAGAGTAAACACACAGCCTATCTAATA 360
 361 TACAATGATGGAGTCCCTAAACCTGTGAATTAAGTAAAGATATGGACACACCAAGGT 420
 361 TACAATGATGGAGTCCCTAAACCTGTGAATTAAGTAAAGATATGGACACACCAAGGT 420
 421 TTACTGCTGTGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 421 TTACTGCTGTGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCT 480
 481 CCAATTCGGGAAGAGTATGATTAATCCACACAGGAGAGGAAATGGACAAAGTGGC 540
 481 CCAATTCGGGAAGAGTATGATTAATCCACACAGGAGAGGAAATGGACAAAGTGGC 540
 541 ATCTGATTAATCTTCAAGTAAACAGATATGAGGCAATAGATTCTCTGCTCTGTCGC 600
 541 ATCTGATTAATCTTCAAGTAAACAGATATGAGGCAATAGATTCTCTGCTCTGTCGC 600
 601 AACCCCAAGCTATAGTCT 660
 601 AACCCCAAGCTATAGTCT 660
 661 CCCAGCTGTGACACAGGAGGAGTCTATCAGAGATTTCTGTCAGGCTCTCTCAGCATT 720
 661 CCCAGCTGTGACACAGGAGGAGTCTATCAGAGATTTCTGTCAGGCTCTCTCAGCATT 720
 721 CAGTGGGCCAGGAGCAAAAAATTCCTCCATTTTGGCAAGTGGATTCCTTTCTTGACGAC 780

Db 721 CAGTGGGCCAGGAGCAAAAAATTCCTCCATTTTGGCAAGTGGATTCCTTTCTTGACGAC 780
 QY 781 ATCTTTGCAGCTGGATGGCTCAACGGCTCAAGACACACACTTTTAAACAGAAACCTGGCAG 840
 Db 781 ATCTTTGCAGCTGGATGGCTCAACGGCTCAAGACACACACTTTTAAACAGAAACCTGGCAG 840
 QY 841 CGAAAAAGACAAGAGCTCTCTCAAACTGCTCCCTTACCCTGCTCTCAATATATAAAA 900
 Db 841 CGAAAAAGACAAGAGCTCTCTCAAACTGCTCCCTTACCCTGCTCTCAATATATAAAA 900
 QY 901 GCAATTAATATATACAGACACTCTTTTTCAGTCTTATCAAGATACAGCAAGTGTGT 960
 Db 901 GCAATTAATATATACAGACACTCTTTTTCAGTCTTATCAAGATACAGCAAGTGTGT 960
 QY 961 ATTTCCCAAAAGGGCCACCAAAAATCGCTGGACATGATTGGAGACCTTAAATCGGAGTCCA 1020
 Db 961 ATTTCCCAAAAGGGCCACCAAAAATCGCTGGACATGATTGGAGACCTTAAATCGGAGTCCA 1020
 QY 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGAAATTTGCAAAATTTACCAAGCA 1080
 Db 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGAAATTTGCAAAATTTACCAAGCA 1080
 QY 1081 TTTTCAAGGATTTAGTATTTATATGAAAGCTGTAAAGTAACTTGGTAAAGGACACAGGT 1140
 Db 1081 TTTTCAAGGATTTAGTATTTATATGAAAGCTGTAAAGTAACTTGGTAAAGGACACAGGT 1140

RESULT 5
 AAF60716
 ID AAF60716 standard; cDNA; 1086 BP.
 AC AAF60716;
 XX
 DT 03-MAY-2001 (first entry)
 XX Human DNase coding sequence.
 DE Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
 XX infectious disease; ss.
 KW Homo sapiens.
 OS
 XX WO200112793-A1.
 XX 22-FEB-2001.
 XX
 PF 01-MAY-2000; 2000WO-JP002893.
 PR 17-AUG-1999; 99JP-00230870.
 XX (TANU/) TANUMA S.
 PA Tanuma S, Shiokawa D;
 PI
 XX WPI; 2001-218348/22.
 DR P-PSDB; AAB72417.
 XX
 PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful
 XX in remedies for cystic fibrosis and for treatment of infectious diseases.
 PS Claim 16; Page 53-55; 61pp; Japanese.
 XX The present sequence is the coding sequence of a human deoxyribonuclease
 CC (DLAD), which is an endonuclease. DLAD is capable of divalent cation-
 CC independent cleavage of DNA under acidic conditions. DLAD can be used as
 CC a substitute for DNase I in treating cystic fibrosis, and is useful in
 CC the prevention and treatment of infectious diseases
 XX
 SQ Sequence 1086 BP; 339 A; 250 C; 228 G; 269 T; 0 U; 0 Other;

Query Match 85.1%; Score 1079.6; DB 4; Length 1086;

	Best Local Similarity	99.6%; Pred No. 0;	Mismatches	4; Indels	0; Gaps	0;
	Matches 1082; Conservative	0;				
QY	34	ATGAAACAGAAAAATGATGGCAAGACTCCTAAGAACATCCCTTTGTCTTCTTCCTCGGC	93			
Dd	1	ATGAAACAGAAAAATGATGGCAAGACTGCTAAGAACATCCCTTTGTCTTCTTCCTCGGC	60			
QY	94	CTCTTTGGGGTGCTGGGGCAGCAACAATTTCATGCCGAATGAAGAAGGAGAAAGCTGTG	153			
Dd	61	CTCTTTGGGGTGCTGGGGCAGCAACAATTTCATGCCGAATGAAGAAGGAGAAAGCTGTG	120			
QY	154	GACTGGTTTTACTTTTTTAAGTTTACCCTAAAAGACAAAACAAAGGAAAGTGGAGAGACTGGG	213			
Dd	121	GACTGGTTTTACTTTTTTAAGTTTACCCTAAAAGACAAAACAAAGGAAAGTGGAGAGACTGGG	180			
QY	214	TTAGAGTACCTGTACTTAGA CTCTACAACTAGAAGCTGGAGAGAGTGGACCACTAATG	273			
Dd	181	TTAGAGTACCTGTACTTAGA CTCTACAACTAGAAGCTGGAGAGAGTGGACCACTAATG	240			
QY	274	AATCACACCAAGAGTGTTCCTGGGAAGGACATTACAAACAGCTATATGAAGCATATGCGCTCT	333			
Dd	241	AATGACACCAAGAGTGTTCCTGGGAAGGACATTACAAACAGCTATATGAAGCATATGCGCTCT	300			
QY	334	AAGAGTAAACAACACAGCCTATCTTAATAFCAANTGAGGTCCCTAAACCTGTGAAATTAC	393			
Dd	301	AAGAGTAAACAACACAGCCTATCTTAATAFCAANTGAGGTCCCTAAACCTGTGAAATTAC	360			
QY	394	AGTAAAGATATGGACACACAAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGGTTCCTGG	453			
Dd	361	AGCAAAAGATATGGACACACAAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGGTTCCTGG	420			
QY	454	CTGATTTCATTCCATCCCTCAGTTTCCTCCAATTCGGAAGAGGCTATGATTATCCACCC	513			
Dd	421	CTGATTTCATTCCATCCCTCAGTTTCCTCCAATTCGGAAGAGGCTATGATTATCCACCC	480			
QY	514	ACAGGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTCAAGTACAAACAGTATGAG	573			
Dd	481	ACAGGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTCAAGTACAAACAGTATGAG	540			
QY	574	GCAATAGATTCTCAGCTCTTGTCTGTCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	633			
Dd	541	GCAATAGATTCTCAGCTCTTGTCTGTCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	600			
QY	634	ACCTTTTCCACGAGAGCTCAATTCACATGCCCCCAGCTGTGCACAGGGCAGCTCATCAGAG	693			
Dd	601	ACCTTTTCCACGAGAGCTCAATTCACATGCCCCCAGCTGTGCACAGGGCAGCTCATCAGAG	660			
QY	694	ATTCTCGGAGGCTCTCACCACATTCAGTCTGGCCAGGGACAAAATTCCTCCATTTT	753			
Dd	661	ATTCTCGGAGGCTCTCACCACATTCAGTCTGGCCAGGGACAAAATTCCTCCATTTT	720			
QY	754	GCAAGTCGGATTCTTTTCTTGACGACATCTTTGACGCTGGATGGCTCAAAGGCTGAAG	813			
Dd	721	GCAAGTCGGATTCTTTTCTTGATGSCATCTTTGACGCTGGATGGCTCAAAGGCTGAAG	780			
QY	814	ACACACTTCTTTAACAGAAA CTTGGCAGCGGAAAAGA CAAGAGCTTCCTTCAAACCTGCTC	873			
Dd	781	ACACACTTCTTTAACAGAAA CTTGGCAGCGGAAAAGA CAAGAGCTTCCTTCAAACCTGCTC	840			
QY	874	CCTTCCTTCCATGCTCTACAAATAAAGCAATTAATAATTCACGACACTCTTATTTTTCAGT	933			
Dd	841	CCTTCCTTCCATGCTCTACAAATAAAGCAATTAATAATTCACGACACTCTTATTTTTCAGT	900			
QY	934	TCTTATCAAGATCACGCCAGTGGTGTATTTCCAAAAGGGCACCAAAAATTCGCTGGACA	993			
Dd	901	TCTTATCAAGATCACGCCAGTGGTGTATTTCCAAAAGGGCACCAAAAATTCGCTGGACA	960			
QY	994	TGTTATTTGGAGACCTTAATTCGAGTCCACACCAAGCCTTCAGAGTGGAGATTTCATTTGT	1053			
Dd	961	TGTTATTTGGAGACCTTAATTCGAGTCCACCAAGCCTTCAGAGTGGAGATTTCATTTGT	1020			
QY	1054	ACCCAGAAATTGGCAAAATTTTACCAGCAATTTCAAGGATTAGTATTTATCTATGAAGCTGT	1111			

Db	1021	ACCAGAAATGGCAAAATTTACCAAGCATTTTCAAGGATTAGTAGTTATTATATGAAAGCTGT	108
QY	1114	AACTAA 1119	
Db	1081	AACTAA 1086	
RESULT 6			
AAF60715			
ID	AAF60715	standard; cDNA; 1652 BP.	
XX	AAF60715;		
AC			
XX			
DT	03-MAY-2001	(first entry)	
XX			
DE		Murine DNase coding sequence.	
XX			
XX		Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;	
KW		infectious disease; ss.	
XX			
OS		Mus musculus.	
XX			
PN	WO200112793-A1.		
XX			
PD	22-FEB-2001.		
XX			
EF	01-MAY-2000; 2000WO-JP002893.		
XX			
PR	17-AUG-1999; 99JP-00230870.		
XX			
PA	(TANU/) TANUMA S.		
XX			
PI	Tanuma S, Shiohara D;		
XX			
DR	WPI; 2001-218348/22.		
DR	P-PSDB; AAB72416.		
PT		Acidic deoxyribonuclease capable of divalent cation-independent cleavage	
PT		of DNA under acidic even neutral pH and not inhibited by G-actin, useful	
PT		in remedies for cystic fibrosis and for treatment of infectious diseases.	
XX			
PS	Claim 14; Page 49-51; 61pp; Japanese.		
XX			
CC		The present sequence is the coding sequence of a murine deoxyribonuclease	
CC		(DLAD), which is an endonuclease. DLAD is capable of divalent cation-	
CC		independent cleavage of DNA under acidic conditions. DLAD can be used as	
CC		a substitute for DNase I in treating cystic fibrosis, and is useful in	
CC		the prevention and treatment of infectious diseases	
XX			
SQ	Sequence 1652 BP; 494 A; 410 C; 339 G; 409 T; 0 U; 0 Other;		
Query Match	53.1%;	Score 673.8; DB 4; Length 1652;	
Best Local Similarity	75.1%;	Pred. No. 1.3e-194;	
Matches 884; Conservative	0; Mismatches 282; Indels 11; Gaps 3		
QY	37	AAACAGAAATGATGGCAGACTGCTAAGAACATCCTTTGCTTGTCTTCTTCTTGGCCCTC 96	
Db	204	AGATAGAAATGACAGCAAGCCTCTAAGACAGTTCTTTTGTCTTCTTGTGCCCTC 263	
QY	97	TTTGGGGTCTGGGGCAGCAACAATTCATGACAGAAATCAAGAGGAAAGCTGTGGAC 156	
Db	264	TCTGGGGTCTGGGGACCAAGAAATCTCATGAGAAATGAATATGTTGAAGCTGTGGAC 323	
QY	157	TGCTTTTACTTTTATTAAGTTACTCTAAAGACAAAACAAGAAAGTGGAGAGACTGGGTTA 216	
Db	324	TGCTTTTACTTTTATTAAGTTATCCCAAAGGACTAGCAAGCAAGTCAAGAGGGGGGCTG 383	
QY	217	GAGTACCTTACTAGACTCTCAACTAGAGAGCTGGAGGAGAGCTGAGCACTAATCAAT 276	
Db	384	CAGTACCTTACTTGTGACTCCCAAGACAAACCTGGAACAGAGCCTTCTACCTGATTAAC 443	
QY	277	GACACCAAGAGTGTGTTGGAGGACATTACACAGCTATATCAAGCATATGCTCTTAAG 330	

Db 444 AGCACCAGGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACACATAATTTCCACG 503
 QY 337 AGTAAACACACAGCGCTACTTAATATACAATGATGAGTCCCTAAACCTGTGAATACAGT 396
 Db 504 AATGACA---CAGCCTATCTAATATACAAACGATGGTCTCCCTGGATCTGTGAATACAGC 560
 QY 397 AGAAGTATGGACACACAAAGGTTTACTGTGTGGAAACAGAGTTCAGGGTCTTGGCTG 456
 Db 561 AGACAGTATGGACATGCGCAAGGTCGTGTGTATGGAACAGAACGCGAGGGTCTTGGCTG 620
 QY 457 ATTCAATTCATCCTCAGATTTCTCCCAATTCGGAAGAGGCTATGATTAATCCACCCACA 516
 Db 621 ATACACTCTGTTCCAGATTTCCCGAGTTC-----ATGGCTATGATGATCCCAACCTCG 674
 QY 517 GGGAGAGCAATGGACAAAGTGGCATCTGCATTAACCTTTCAAGTACACACCACTGATGAGGCA 576
 Db 675 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTTCGGATACACCGCAGTTTGAGAA 734
 QY 577 ATAGATTTCTCAGCTCTTGGTCTGCAACCCCAACGCTATATAGTGTCTCCATCCCGACCC 636
 Db 735 ATAGATTTCTCAGCTCTTGGTCTTACACCAACCAATCTACAGCTGCTTCAATCCCAAGCA 794
 QY 637 TTTTCCAGGAGTCTCATTCATGCCCCCAGCTGTGCACACGAGGCGCAGCTCATCAGAGATT 696
 Db 795 TTTTCACTGGAACTTATCTACATGCCCGGATGTGTGCCAACTCCAGTTTCTTTAAAGATC 854
 QY 697 CTGGCAGGCTCTCCACCACTCTAGTGGCGCCAGGACAAAATTTCTTCCATTTTGA 756
 Db 855 CTTGTCCGGTACCTGCTGAACCTGCACTCAGCCAGGGTCTAAACTTGTCTCCATTTTGA 914
 QY 757 AAGTCGGATCTTTCTTACGACATCTTTGACGCTGTGGAGGCTCAACGGCTGGAAGACA 816
 Db 915 AAATCAAGTTTTTATCTGATGACATCTTTTACAGGATGATAGCTCAAAAGTTTGAAGACA 974
 QY 817 CACTTGTGTTAAACAACTGGCAGCGAAAGACAGAGCTTCTTCAACTGCTCCCTT 876
 Db 975 CATTTGTTAGCAACAACTGGCAGAAAAGAACAAAGAGCTTCTTCAAACTGTTCCTG 1034
 QY 877 CTTTACCACTGTCTAATATAAAGCAATTAATATACGACACTCTTATTTTCACTTCT 936
 Db 1035 CTTTACCACTGTCTAACAATCAAGTCCATTGGGTAACCTTCCAAGTCTTACTTCACTTCT 1094
 QY 937 TATCAAGTACGCGCAAGTGTGTATTTCCCAAAAGGGCCACAAATCGCTGGACATGT 996
 Db 1095 CCCCAGGACCATTTCCAAATGGTGTGTTCCTAATAAGGGCTCCGCAATCGCTGGACCTGC 1154
 QY 997 ATTGGAGACCTAAATCGGAGTCCACCAAGCCTTCCAGAGTGGAGGATTCATTTGTACC 1056
 Db 1155 ATTGGAGACCTAAATCGAGCCTACACCAAGCCTTAAAGAGTGGAGGATTCATCTGTACA 1214
 QY 1057 CAGAATGGCAATTTACCAAGCATTTCAAGGATTAGTATTAATATCTATGAAAGCTGTAA 1116
 Db 1215 AAGAATCACTACATTTACAGGCAATTCATAAATTAATATCTCCGTATGGGTCTGTAA 1274
 QY 1117 TAACTTGGTGAAGGACACAGGTAATCATTTGAAACCTTGAACATGGGTCTTCTTCC 1176
 Db 1275 TAACTGGTGAAGGACACACCTCTCTCTTCTTGAACACATGGCACTGGCAATCTGCG 1334
 QY 1177 ATTACAC--TTCTTTATATTTTAAAGGCTGTGAAT 1211
 Db 1335 CTTGATCTGTCTCCATTAATTTCAAGGCTTCTGAGT 1371

RESULT 7

AAD21288
 ID AAD21288 standard; cDNA; 1224 BP.
 XX
 AC AAD21288;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Mouse deoxyribonuclease (DNase) II beta cDNA.
 XX

Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
 cystic fibrosis; pulmonary; ss.

Mus sp.

Key Location/Qualifiers
 CDS 40..1104
 /*tag= a
 /product= "Mouse deoxyribonuclease (DNase) II beta
 protein"

WO200175082-A1.

11-OCT-2001.

02-APR-2001; 2001WO-US010635.

03-APR-2000; 2000US-00541840.

19-MAY-2000; 2000US-00574942.

(DART-) DARTMOUTH COLLEGE.

Eastman AR, Krieser RJ;

WPI; 2001-662972/76.

P-PSDB; AAE13014.

New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
 DNA present in the mucous plugs in the lungs of cystic fibrosis patients.

Claim 2; Page 13-14; 21pp; English.

The invention relates to deoxyribonuclease (DNase) II beta proteins and
 their corresponding cDNAs. The DNase II beta may be useful to digest DNA
 in the mucous plugs in lungs of cystic fibrosis patients and so reduce
 their viscosity. The present sequence is mouse DNase II beta cDNA

Sequence 1224 BP; 351 A; 308 C; 256 G; 309 T; 0 U; 0 Other;

Query Match 53.1%; Score 672.8; DB 5; Length 1224;
 Best Local Similarity 75.0%; Pred. No. 2.3e-194;
 Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 37 AAACAGAAATGATGCGAGAGCTGTAAGAAACATCTTTGCTTTGCTTCTTCTTGGCCTC 96
 Db 31 AGATAGAAATGACAGCAAGGCTCTAAGAACAGTCTTTCTTCTTCTTCTTCTTCTTCT 90
 QY 97 TTTGGGTGCTGGGGGCGAGCAACATTTTCATGCAAGATGAAGAGGAAAGCTGTGGAC 156
 Db 91 TCTGGGTCTTGGGGGACACCAAGAAATCTCATGCAAGAAATGAATATGTTGAAGCTGTGGAC 150
 QY 157 TGGTTTACTTTTTATAGTTTACTTAAAAAGACAAACAGGAAAGTGGAGAGACTGGTTA 216
 Db 151 TGGTTTACTTTTTATAGTTTACTTAAAAAGGACTAGCAAGGCAAGTGAAGAGCGGGCTG 210
 QY 217 GAGTACCTGTACTAGCTCTACAACTAGAGCTGGAGGAAAGAGTGAACAATTAATGAAT 276
 Db 211 CAGTACCTGTACTTGGACTCCACCAAGCAACCTTGAACAGAGCTCTACTGATTAAC 270
 QY 277 GACACCAAGAGTGTGGGAGGACATTAACAAGCTATATGAAGCATATGCTCTTAAG 336
 Db 271 AGCACCAGGAGTGTCTGGGAGGAGCCTTACAGATCTGTATGACACACATAATTCACG 330
 QY 337 AGTAAACACACAGCCTATCTAATATACATGATGAGTCCCTTAACCTGTGAATTAAGT 396
 Db 331 AATG---ACACAGCTATCTAATATACACATGCTGCTCCCTGGATCTGTGAATTAAGC 387
 QY 397 AGAAGTATGGACACACCAAGGTTTACTGTGTGGAAACAGAGTTCAGAGGTTCTGCTG 456
 Db 388 AGACAGTATGGACATGCGCAAGGTCGTGTGTATGGAACAGAGCAGGAGGTTCTGGCTG 447
 QY 457 ATTCAATCCATCCCTCAGTTTCTCTCAATTCGGAAGAGGCTATGATTAATCAACCCACA 516

Fri Oct 15 10:01:04 2004

us-10-790-589-3.rng

Db 448 ATACACTGTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCAACCTCG 501
QY 517 GCGAGACGAAATGGACAAAGTGGCATCTGCATTAACCTTTCAAGTACCAACAGTATGAGGCA 576
Db 502 GCGAGCGGATAGGACAAACCGGCATCTGCATCACTTTCGGATACACAGCCAGTTTGAGGAA 561
QY 577 ATAGATTCAGCTCTTGTGTTGCTCCAAACCCCAAGCTGTATAGCTGTCTCCATCCAGCCACC 636
Db 562 ATAGATTTTCAGCTCTTGTGTTGCTCCAAACCCCAAGCTGTATAGCTGTCTCCATCCAGCCACC 621
QY 637 TTTTCCACGAGCTCATTTCAATGCCCCCAGCTGTGACACAGGCCAGCTCATCAGAGATT 696
Db 622 TTTTCACTGGAACTTATCTACATGCCCCCGGATGTGTCGCACTCCAGTTCCTTAAAGATC 681
QY 697 CTGCGAGGCTCTCCACCACTTTCAGTTCGCGCCCGAGGACAAAATTCCTCCATTTTGCA 756
Db 682 CTTGTCCGGTACCTCGTGAACCTGCACTCAGCCCGAGGCTCTAACTTCGTCCATTTTGCA 741
QY 757 AAGTCGGATCTTTCTTGACGACATCTTTGACGCTTGGACCTGGATGCTCAACGCTCAAGACA 816
Db 742 AAATCAAGTTTATATCTGATGATCATCTTTACAGGATGGATAGCTCAAAAAGTTGAAGACA 801
QY 817 CACTGTGTAACAGAACTGTCAGCGGAAAGACAAAGAGCTTCTTCAAACTGCTCCCTT 876
Db 802 CATTTGTTAGCAAACTGTCAGGAAAGAGAAAGAGAGCTTCTTCAAACTGCTCCCTG 861
QY 877 CTTTACCATGCTCTCAATATAAAGCAATTAATATACGACACTCTTATTTTCAGTTCT 936
Db 862 CTTTACCATGCTCTCAATATAAAGCAATTAATATACGACACTCTTATTTTCAGTTCT 921
QY 937 TATCAAGATCAGCGCAAGTGTGTATTTCCCAAAGGCGCACCAAAATCGCTGACATGT 996
Db 922 CGCCAAAGACCATTCCAAATGTTGTGTTTCCATTAAGGGCTCCGCAATTCGCTGACCTGC 981
QY 997 ATTGAGACCTTAATCGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGACC 1056
Db 982 ATTGAGACCTTAATCGAGTCCACCAAGCTTTCAGAAAGTGGAGGATTCATTTGACA 1041
QY 1057 CAGAAATGGCAAAATTTACCAAGCATTTCAAGGATTAATATATGATTAATGAAGCTGTAA 1116
Db 1042 AAGAATCACTACATTTACCGAGCATTTTCAAAATATATCTCCGTTATGGGTTCTGTAA 1101
QY 1117 TAAACTTGGTGAAGGACACAGTACTATCATTTGAAACCTTGCAATGGTCTTCTTCC 1176
Db 1102 TAAACTCGTGAAGGCGCACACCTCTGCTCTTGTGAAACACTGGCACTGGAACTCTCGC 1161
QY 1177 ATTAACCTTCTTTATATTTA 1198
Db 1162 CTGGATCTGTCTCCATAATA 1183

RESULT 8
ACH19937
ID ACH19937 standard; cDNA; 480 BP.
AC ACH19937;
XX
XX
XX 13-OCT-2003 (first entry)
XX Human adult lung cDNA #940.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX

XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX New polynucleotide sequences obtained from various cDNA libraries. useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 7149; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: the sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 480 BP; 134 A; 120 C; 106 G; 118 T; 0 U; 2 Other;

Query Match 24.5%; Score 310.8; DB 8; Length 480;
Best Local Similarity 99.4%; Pred. No. 4.9e-84;
Matches 312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 335 AGAGTAAACACACAGCCTATCTAATATACATGATGAGTGCCCTAAACCTGTGAATTACA 394
Db 166 AGAGTAAACACACAGCCTATCTAATATACATGATGAGTGCCCTAAACCTGTGAATTACA 225
QY 395 GTAGAAAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGTTCTGCGC 454
Db 226 GCAGAAAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGTTCTGCGC 285
QY 455 TGATTTCATTCCTCCTCAGTTTCTCCAAATTCGGAAGAGGCTATGATTATCCACCA 514
Db 286 TGATTTCATTCCTCCTCAGTTTCTCCAAATTCGGAAGAGGCTATGATTATCCACCA 345
QY 515 CAGGAGACGAAATGGACAAAGTGGCATCTGCATTAACCTTCAAGTACACAGGATGAGG 574
Db 346 CAGGAGACGAAATGGACAAAGTGGCATCTGCATTAACCTTCAAGTACACAGGATGAGG 405
QY 575 CAATAGATTTCTCAGCTCTTGGTCTGCAACCCCAAGCTGTATAGCTGTCTCCATCCAGGCA 634
Db 406 CAATAGATTTCTCAGCTCTTGGTCTGCAACCCCAAGCTGTATAGCTGTCTCCATCCAGGCA 465
QY 635 CCTTTTACCAGGAG 648
Db 466 CCTTTTACCAGGAG 479

RESULT 9
ADE87477/c
ID ADE87477 standard; DNA; 266145 BP.
XX AC ADE87477;
XX

OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FF	68..1150
CDS	/*tag= a
FT	/product= "deoxyribonuclease (DNase II)"
FT	
XX	
PN	W09816659-A1.
XX	
PD	23-APR-1998.
XX	
PF	09-OCT-1997; 97WO-US018262.
XX	
PR	15-OCT-1996; 96US-0028539P.
XX	
PA	(DART-) DARTMOUTH COLLEGE.
XX	
PI	Eastman A, Krieser R;
PI	
DR	WPI; 1998-251301/22.
XX	
DR	P-PSDB; AAW37920.
XX	
XX	Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
PT	e.g. to induce apoptosis in tumour cells, identify agents modulating
PT	apoptosis and digest DNA in sputum of cystic fibrosis patients.
XX	
PS	Claim 2; Page 15-16; 29pp; English.
XX	
CC	The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme
CC	is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis
CC	patients to reduce sputum viscosity. It (or fragments) are also useful to
CC	produce antibodies, e.g. to study DNase II expression in cells. DNase II
CC	has recently been linked with DNA fragmentation in the early stages of
CC	controlled cell death (apoptosis), a process critical to homeostasis
CC	during, e.g. metamorphosis or cell turnover. Too much cell death can lead
CC	to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
CC	whilst too little can lead to cancer or autoimmune diseases. The
CC	antibodies raised against it can therefore be used to diagnose apoptotic
CC	stages in selected cells, by contacting cells with the antibody,
CC	detecting binding of the antibody with DNase II and determining DNase II
CC	levels. Such diagnosis is useful to evaluate the efficacy of therapeutic
CC	agents, e.g. anticancer agents to promote apoptosis in cells. The
CC	antibodies can also be used to identify cells susceptible to premature
CC	death. The cDNA is useful to identify agents modulating apoptosis in
CC	cells, by treating cells with an agent, transfecting cells with cDNA and
CC	monitoring apoptosis compared with untreated cells. Inhibitors identified
CC	may be useful in preventing diseases relating to enhanced chromosomal re-
CC	arrangement. Vectors comprising the cDNA can be used to induce apoptosis
CC	in selected cells, e.g. tumour cells or cells involved in autoimmune
CC	disorders. Antisense oligonucleotides can be administered to cells to
CC	inhibit DNase II expression to reduce chromosome instability associated
CC	with cancer
XX	
SO	Sequence 1915 BP: 462 A; 533 C; 530 G; 390 T; 0 U; 0 Other;

	Query Match	6.6%;	Score 84;	DB 2;	Length 1915;
	Best Local Similarity	47.5%;	Pred. No. 1.9e-14;		
	Matches 471;	Conservative	0;	Mismatches 485;	Indels 36; Gaps 6;
QY	142	GGGAAGCTGTGGACGTGGTTACTTTTATAAGTT---ACCTAAAAGACAAAACAAGAA	198		
Db	137	GGGCGAGCTGTAGACTGGTTCGTCTACAAGCTGCCAGCTCTAGAGGTCCTGGGGAG	196		
QY	199	AGTGAGAGACATGGGTTTAGAGTACCTGTCTACCTAGACTCTACAATAGACAGCTGGAGGAAG	258		
Db	197	GCGGCGCGAGAGGGCGTGCAGTACAAAGTATCTGGACAGAGACTCCGGAGGTTGGCGGAC	256		
QY	259	AGTGAGCACTAATGAATGACCAAGAGTGTTTTGGGAAGGACATTCACACAGCTATAT	318		
Db	257	GCAGGGCACTCATCAACAGCCGGAGGGGGCGCTGGGCCGAAGCCTGACCGCGCTGAC	316		
QY	319	GAAGCATATGCTCTAAGTAGTACAACACAGCGCTATCTAATATACAATGATGGAGTCCT	378		

FT CDS /note= "Region indicated in specification"
 FT 86..1168
 FT /*tag= b
 FT /product= "DNase II"
 FT sig_peptide 86..133
 FT /*tag= c
 FT mat_peptide 134..1165
 FT /*tag= d
 FT /product= "DNase II"

XX WO9740134-A2.
 XX 30-OCT-1997.
 XX 23-APR-1997; 97WO-US006664.
 XX 25-APR-1996; 96US-00639294.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Baron WF;
 XX WPI; 1997-535820/49.
 XX P-PSDB; AAW31495.
 XX Human DNase II and related nucleic acids - useful in protein production
 XX e.g. for therapeutic use to treat systemic lupus erythematosus and
 XX pulmonary diseases e.g. cystic fibrosis.
 XX Claim 1; Fig 1; 30pp; English.
 XX This sequence encodes a novel human deoxyribonuclease, DNase II. This
 XX protein is useful to reduce the viscoelasticity of DNA-containing
 XX material, e.g. mucus, and used to treat patients with pulmonary diseases
 XX or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.
 XX It may also improve antibiotic efficacy in the treatment of abscesses,
 XX infected lesions etc., provide treatment in non-infected conditions in
 XX which there is an accumulation of cellular DNA debris, e.g.
 XX pyelonephritis and be used to degrade DNA in biological samples or in
 XX diagnostic assays. The nucleic acids and vectors can be used for in vivo
 XX or ex vivo gene therapy, and antisense oligonucleotides can be produced
 XX from the nucleic acids which can bind to and prevent expression of
 XX nucleic acid within cells. The nucleic acids (or a portion) can also be
 XX used for hybridisation assays for nucleic acids encoding human DNase II
 XX in a sample, or to identify and isolate nucleic acids sharing substantial
 XX sequence identity (e.g. encoding naturally-occurring allelic variants of
 XX human DNase II). The antibodies can be used to detect and measure human
 XX DNase II in tissues or clinical samples, and in the purification of human
 XX DNase II. The nucleic acids enable production of human DNase II by
 XX recombinant DNA methods in quantities sufficient for clinical use, not
 XX previously possible
 XX Sequence 1575 BP; 358 A; 465 C; 450 G; 302 T; 0 U; 0 Other;
 XX
 XX Query Match 6.5%; Score 82.4; DB 2; Length 1575;
 XX Best Local Similarity 47.4%; Pred. NO. 5.3e-14;
 XX Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;
 QY 142 GGGAAAGCTGGAGCTGGTTTACTTTTATAGTT---ACCTAAGACAAACAGGAA 198
 Db |||||
 QY 155 GGGCAGCTGTAGACTGGTTCTGTGTCTACAGCTGCCAGCTCTTAGAGGGTCCGGGAG 214
 Db |||||
 QY 199 AGTGGAGAGCTGGTTAGAGTACTGTACCTAGACTCTACACTAGAGCTGGAGGAG 258
 Db |||||
 QY 215 GCGGCGCAGAGAGGGCTCAGTACAGTATCTGACGAGAGCTCGGAGGCTGGCGGAC 274
 QY 259 AGTACAGCACTAATGATGACACCAAGAGTGTTTTGGGAGGACATTACACAGCTATAT 318
 Db |||||
 QY 275 GGCAGGGCACTCATCAACAGCCGAGGGGCGCGTGGGCCGAGCCCTGACCTGTAC 334
 QY 319 GAAGCATATGCTCTAAGAGTAAACACACAGCCCTATCTAATATACATGATGAGTCCCT 378
 Db |||||
 QY 335 CGG-----AGCAACACGACGCTGCCTTCTGCTCTACATGACCAACCGCCT 385
 Db |||||

QY 379 AAACC-----TGTGAATTACAGTAGAAGTATGACACACAAAGGTTTACTGCTGG 432
 Db |||||
 QY 386 CAACCCAGCAAGGCTCAGGACTCTTCATGCTGGGACACGAAGGGTGTCTGCTCTCTT 445
 Db |||||
 QY 433 AACAGAGTTTCAAGGGTCTGCTGATTCATTCCATTCCTCAGTTTCTTCCAAATTCGGAA 492
 Db |||||
 QY 446 GACCACGATGGGGCTTCTGGCTGTCCACAGTGTACCTAACTTCCCTCCACCGGCTCC 505
 Db |||||
 QY 493 ---GAAGGCTATGATTATCCACCCACAGGAGAGCAAAATGACAAAGTGGCATCTGATA 549
 Db |||||
 QY 506 TCTGCTCATACAGCTGGCTCTATAGCGCTGTACCTACGGGCGAGACCTGCTCTGTGTG 565
 Db |||||
 QY 550 ACTTTCAAGTACAAACAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAC 609
 Db |||||
 QY 566 TCTTTTCCCTTCTGCTCAGTTCTGAAATGGCAAGCAGCTGACCTACACTACCTACCTCC 625
 Db |||||
 QY 610 GTCTATAGCTGCTCCATCCAGCCACTTTTCAACAGGAGCTCATTCACATGCCCACTG 669
 Db |||||
 QY 626 GTCTATAACTACCAAGCTTGGAGGGATCTTTGCCAGGAATTCCTGAGTATGGAGATGTG 685
 Db |||||
 QY 670 TGCACACAGGCGCAGCTCATCAGAGATTCTTGGCAGGCTCTCACCACACTTCAGTCGGCC 729
 Db |||||
 QY 686 GTCAAGGGCCACCAAGTACCAAGAACCTTGGAAACAGCAGCATCACACTCACATCCACG 745
 Db |||||
 QY 730 CAGGGAACAAATTCCTTCCATTTTGAAGTCCGATTCTTTTCTTGACGACATCTTTGCA 789
 Db |||||
 QY 746 GCGGGGCTGTTTTCAGAGCTTTGCCAAGTTTCCAGCAATTTGGAGATGACCTGTACTCC 805
 Db |||||
 QY 790 GCCTGGATGCTCAACGGCTGAAGACACACTTGTAAACAGAACTGGGACGCGAAAAAGA 849
 Db |||||
 QY 806 GGTGTTGGCAGCAGCCCTTGGTACCAACTGCGAGTCCAGTCTGGCACAACACTGTA 865
 Db |||||
 QY 850 CAAGAGCTTCTTCAAACTGCTCCCTTCTTACCATGTCTACAATATAAAGCAATATAA 909
 Db |||||
 QY 866 GGCACTCTCCCTCTAACTGCTCGGATCTGGCAGGTTCTGAATGTGAACCATGATGCT 925
 Db |||||
 QY 910 TTATACGACA-----CTCTTATTTCAGTTCTTATCAAGATCAGCCCAAGTGTGTTAT 963
 Db |||||
 QY 926 TTCCCTGGACCAAGCGCGCCAAAGTTTCAACAGCAGAGGACCACTCCAAATTTGGTGGT 985
 Db |||||
 QY 964 TCCCAAAAGGACCAAAATCGCTGACATGTTTGGAGACCTAAATCGGAGTCCACAC 1023
 Db |||||
 QY 986 TCCCAAAAGGAGGCC-----CTGGACCTCGTGGTGGTACATGAATCGGAACCAAGGA 1036
 Db |||||
 QY 1024 CAAGCTTTCAGAAAGTGGAGGATTCATTTGTACCCAGAAATTCGCAATTTACCAAGATT 1083
 Db |||||
 QY 1037 GAGGAGCAACGGGTTGGGGCAGACTGTGCCCCAGCTGCCAGCCCTCTGGAAAGCCTTC 1096
 Db |||||
 QY 1084 CAAGGATTAGTATTATATATATGAAAGCTGTAA 1115
 Db |||||
 QY 1097 CAGCGCTGTGAAGAACTTACCAGCCCTGTAA 1128
 Db |||||

RESULT 12
 AAV29138
 ID AAV29138 standard; cDNA; 927 BP.
 XX AC AAV29138;
 XX DT 11-SEP-1998 (first entry)
 XX DE Bos taurus DNase II gene.
 XX KW DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
 XX KW controlled cell death; apoptosis; metaplasia; cell turnover;
 XX KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
 XX KW autoimmune disease; diagnosis; anticancer agents;
 XX KW enhanced chromosomal rearrangement; chromosome instability; ss.
 XX OS Bos taurus.
 XX FH Key Location/Qualifiers

CDS	1..831	
FT	/tag= a	
FT	/product= "deoxyribonuclease II (DNase II)"	
XX		
PX	WO9816659-A1.	
PN		
PP	23-APR-1998.	
XX		
XX	09-OCT-1997;	97WO-US018262.
XX		
PR	15-OCT-1996;	96US-0028539P.
PA	(DART-) DARTMOUTH COLLEGE.	
XX		
PI	Eastman A, Krieser R;	
DR	WPI; 1998-251301/22.	
DR	P-PSDB; AA037921.	
XX		
PT	Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,	
PT	e.g. to induce apoptosis in tumour cells, identify agents modulating	
PT	apoptosis and digest DNA in sputum of cystic fibrosis patients.	
XX		
PS	Claim 2; Page 17; 29pp; English.	
XX		
CC	The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme	
CC	is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis	
CC	patients to reduce sputum viscosity. It (or fragments) are also useful to	
CC	produce antibodies, e.g. to study DNase II expression in cells. DNase II	
CC	has recently been linked with DNA fragmentation in the early stages of	
CC	controlled cell death (apoptosis), a process critical to homeostasis	
CC	during, e.g. metamorphosis or cell turnover. Too much cell death can lead	
CC	to neuro-degeneration and acquired immune deficiency syndrome (AIDS),	
CC	whilst too little can lead to cancer or autoimmune diseases. The	
CC	antibodies raised against it can therefore be used to diagnose apoptotic	
CC	stages in selected cells, by contacting cells with the antibody,	
CC	detecting binding of the antibody with DNase II and determining DNase II	
CC	levels. Such diagnosis is useful to evaluate the efficacy of therapeutic	
CC	agents, e.g. anticancer agents to promote apoptosis in cells. The	
CC	antibodies can also be used to identify cells susceptible to premature	
CC	death. The cDNA is useful to identify agents modulating apoptosis in	
CC	cells, by treating cells with an agent, transfecting cells with cDNA and	
CC	monitoring apoptosis compared with untreated cells. Inhibitors identified	
CC	may be useful in preventing diseases relating to enhanced chromosomal re-	
CC	arrangement. Vectors comprising the cDNA can be used to induce apoptosis	
CC	in selected cells, e.g. tumour cells or cells involved in autoimmune	
CC	disorders. Antisense oligonucleotides can be administered to cells to	
CC	inhibit DNase II expression to reduce chromosome instability associated	
CC	with cancer	
XX		
SQ	Sequence 927 BP; 214 A; 258 C; 237 G; 218 T; 0 U; 0 Other;	
	Query Match 5.6%; Score 71; DB 2; Length 927;	
	Best Local Similarity 47.7%; Pred. No. 1.2e-10;	
	Matches 317; Conservative 0; Mismatches 330; Indels 18; Gaps 3;	
QY	405 TGACACACCAAGGTTTACTGTGTGGACAGAGTTTCAGGGGTTCTGGCTGATTCATTTC 464	
DB	75 TGGGCACACGAAGGTTGTGCTCCTGGACAGAAGGGGGGCTTCGTGTGATCCACAG 134	
QY	465 CATCCCTCAGTTTCCTCCAATTCGGGA---AGAAGGCTTGATTAATCACCCACAGGGAG 521	
DB	135 CGTTCCAAACCTCCCTCCACGGTGCTCTCTGCTCGGTACAGCTGGGCTCTCTGTGCCCA 194	
QY	522 ACBAATGACAAAGTGGAATCTGCAATCACTTCAAGTACACAGCATGAGGCAATAGA 581	
DB	195 AAAATATGGGCAGACCCTGATCTGTGTATCTTTCTCTACCCAGTCTCTGGATATCAG 254	
QY	582 TTCTCAGCTCTGGTGTGCAACCCCAAGCTGTATAGCTGTCTCCATCCACGCCACTTCA 641	
DB	255 CAACAGCTGACCTATACCTATCTCCACTGGTATATGAACACAGGCTGGAAGGGGATTTGG 314	
Db	315 CCAGAAATTCCTTACCTGGAGGAGTAGTCAAGGGCCATCACGTTCCGCCAGGACCGTG 374	
QY	702 CAGGCTCCTCCACACACTTCAGTCGGCCCCAGGGACAATAATTCCTCAATTTTCCAAAGTC 761	
DB	375 GAACAGCAGTGTAACTCATCATCAAAAAGAGAGCCACATCTCCAGAGCTTTTGCCTAAAT 434	
QY	762 GGATCTTTTTCTTGACGACATCTTTTGACGCTCGATGGCTCAAGGGCTGGAAGACACACTT 821	
DB	435 TCGAAACTTTTGAGATGACCTGTACTCTGGCTGGCTGGCGGAAGCCCTTGGCAGTACCT 494	
QY	822 GTTAACAGAAACCTGGCAGCGAAGAAAAAGACAAGAGCTTCTTTCAAACTGCTCTCCCTT 881	
DB	495 GCAGGTCCAAATCTTCGCAACGATCTTCTGTGTATCTCTGCTCCCTCAACTGCTCTGGGGCC 554	
QY	882 CCATGTCTACAAATATAAAAGCAATTAATATACGACACTCT-----TATTTTCAGTTC 935	
DB	555 GCATGTATTTGACGTGACTCAGACAGCTTTCCCTGGGCCAGCTGGGCCAGCTTCAATGC 614	
QY	936 TTATCAGATACCGCAAGTGGTGTATTTCCCAGAAAGGGCACCAAAATCCCTCGACATG 995	
DB	615 CACAGAGAACCAATTCAGTGGTGTATACCCCAAAAGGGCC-----CTGGGCTG 665	
QY	996 TATTGAGACCTAAATCGGAGTCCACACCAAGCCCTTCAGAAAGTGGAGATTCATTTGTAC 1055	
DB	666 TGTGGGTGACATGATCGGAACCAAGAGAGAGCAGCAGCGGGTGGGGGCACTCTCTGTGC 725	
QY	1056 CCAGA 1060	
DB	726 CCAGA 730	
RESULT 13		
AEN40474	ID AEN40474 standard; DNA; 60 BP.	
XX	AEN40474;	
AC	AEN40474;	
XX	15-JUL-2002 (first entry)	
DT	Human spliced transcript detection oligonucleotide SEQ ID NO:13222.	
DE	Human; mouse; rat; splice transcript; detection; RNA transcript;	
KW	splice variant; transcriptome; oligonucleotide library; ss.	
XX	Homo sapiens.	
OS	WO200210449-A2.	
XX	07-FEB-2002.	
PN	20-JUL-2001; 2001WO-IB001903.	
PD	28-JUL-2000; 2000US-0221607P.	
XX	02-MAY-2001; 2001US-0287724P.	
XX	(COMP-) COMPUGEN INC.	
PF	Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;	
XX	WPI; 2002-257383/30.	
XX	New oligonucleotide libraries comprising oligonucleotides which	
DR	selectively hybridize to mRNAs transcribed from a transcription unit	
XX	genome, useful for detecting tissue-, pathology-, and developmental-	
PT	specific genes.	
PT	Example 1; SEQ ID NO 13222; 47pp; English.	
XX	The present invention describes oligonucleotide libraries for detecting	
CC	messenger RNAs that populate a (sub-)transcriptome, where the (sub-	
CC)transcriptome comprises messenger RNAs transcribed from multiple	
CC	transcription units that populate a genome. The library comprises several	

XX

PS Example 9; Page 934; 959pp; English.

XX The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA, sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a foetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library

XX Sequence 544 BP; 114 A; 175 C; 133 G; 122 T; 0 U; 0 Other;

Query Match 4.3%; Score 54.4; DB 7; Length 544;
Best Local Similarity 47.5%; Pred. NO. 1.1e-05;
Matches 226; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

QY	400	AAGTATGACACACCAAGGTTTACTGCTGTGGACAGATTCAAGGGTTCTGGCTGATT	459
DB	57	ATGCGTGGGCACACGAGGGTGTCTGCTCTTTGACACGATGGGGGCTTTGGCTGTC	116
QY	460	CATTCCATCCCTCAGTTTCTCCAA---TTCGGAAGAAGGCTATGATTATCCACCCACA	516
DB	117	CACAGTGACUAACTTCTCCCTCCCGGCTCTCTGCTGCATACAGCTGGCTCATAGC	176
QY	517	GGGAGACGAAATGGAACAAAGTGGCATCTGCATACTTTCAAGTACAAACAGTATGAGCA	576
DB	177	GCTCTACTACGGGCAGACCCCTGCTCTGTGTCTTTTCCCTCGCTCAGTTCTCGAAG	236
QY	577	ATAGATTCTCAGCTCTGGTCTGGAACCCCACTCTAGCTGTCTCCATCCAGCCACC	636
DB	237	ATGGGCAAGCAGCTACCTACCTACCCCTGGGCTCTATACTACAGCTGAAGGATC	296
QY	637	TTTCCACGAGGAGCTCATTCACATGCCCGCAGCTGTGCACAGGGCCAGCTCATCAGAGATT	696
DB	297	TTTGCCCAAGAAATCCCGACTTGGAGATGTGGTCAAGGGCCACCACCTAGCCAGAA	356
QY	697	CCTGGCAGGCTCTCACCACACTTCAGTCGGCCAGGGACAAAAATTCCTCATTTTGA	756
DB	357	CCCTGGAACAGCAGCATCACACTCACTCCAGCCGGGCTGTT-TTCCAGAGCTTTGCC	415
QY	757	AAGTCGATCTTTTCTTTCAGCAGCATCTTTGAGCTGTGATGGCTCAACGGCTGAAGACA	816
DB	416	AAGTTACGCAAAATTTGGAGATGACCTGTACTCCGGCTGGTTGGCAGCAGCCCTTGTACC	475
QY	817	CAGTTGTTAACAGAAACCTGGCAGCGAAGAAAGACAGAGCTTCCTTCAAACTGCTC	872

Db 476 AACCTGACGTCAGTTCTGGCACAAAACCTAGGATCCTGCCCTCTACTGCTC 531

Search completed: October 14, 2004, 01:02:06
Job time : 487.369 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:43:42 ; Search time 106.854 Seconds
(without alignments) 6585.417 Million cell updates

Title: US-10-790-589-3

Perfect score: 1268

Sequence: 1 atggggaagtgtcctgctg.....tctctcatgtttaccattta 1268

Scoring table: IDENTITY NIC

```

=====
IDENTIFIER_NUC
Gapop 10.0 , Gapext 1.0
=====

```

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

[illegible]

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:**

```

1: /cgm2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgm2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgm2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgm2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgm2_6/ptodata/2/ina/PCTUS_COMB.seq
6: /cgm2_6/ptodata/2/ina/backfiles1.seq

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1268	100.0	1268	4	US-09-574-942-3		Sequence 3, Appli
2	1079.6	85.1	1086	4	US-09-807-784B-4		Sequence 4, Appli
3	673.8	53.1	1652	4	US-09-807-784B-2		Sequence 2, Appli
4	672.8	53.1	1224	4	US-09-574-942-1		Sequence 1, Appli
5	84	6.6	1915	3	US-09-147-915-1		Sequence 1, Appli
6	82.4	6.5	1575	3	US-08-639-294-1		Sequence 1, Appli
7	82.4	6.5	1575	4	US-09-861-034B-1		Sequence 1, Appli
8	71	5.6	927	3	US-09-147-915-2		Sequence 2, Appli
9	51.8	4.1	7218	1	US-08-232-463-14		Sequence 14, Appl
10	40.2	3.2	505	4	US-09-621-976-15639		Sequence 15639, A
11	39.4	3.1	832	4	US-09-621-976-2813		Sequence 2813, A
12	39	3.1	169938	4	US-09-676-610B-24		Sequence 24, Appl
13	39	3.1	197496	4	US-09-877-177A-10		Sequence 10, Appl
14	34.6	2.7	1509	4	US-09-134-000C-3209		Sequence 3209, Ap
15	34.6	2.7	3321	1	US-08-484-438-5		Sequence 5, Appli
16	34.6	2.7	5484	3	US-09-632-580A-3		Sequence 3, Appli
17	34.6	2.7	5501	1	US-08-484-438-1		Sequence 1, Appli
18	34.4	2.7	364	4	US-09-621-976-17202		Sequence 17202, A
19	34.4	2.7	1830121	4	US-09-557-884-1		Sequence 1, Appli
20	34.4	2.7	1830121	4	US-09-643-990A-1		Sequence 1, Appli
21	34.2	2.7	8779	4	US-08-956-171E-58		Sequence 58, Appl
22	34	2.7	3887	4	US-09-976-594-660		Sequence 660, App
23	33.8	2.7	551	4	US-09-389-681-250		Sequence 250, App
24	33.8	2.7	551	4	US-09-620-405B-250		Sequence 250, App
25	33.8	2.7	551	4	US-09-333-338-250		Sequence 250, App
26	33.8	2.7	551	4	US-09-433-826B-250		Sequence 250, App
27	33.8	2.7	551	4	US-09-604-287A-250		Sequence 250, App

361	TACAAATGATGGAGTCCCTAAACCTTGATTAACAGTAGAAAGTATGGACACACCAAGGT	420
421	TTACTGCTGTGGAAACAGAGTTCAAGGGTCTGGCTGATTCATTCCATCCCTCAGTTTCCT	480
421	TTACTGCTGTGGAAACAGAGTTCAAGGGTCTGGCTGATTCATTCCATCCCTCAGTTTCCT	480
481	CCAAATCCGGAGAGGCTATGATTTATCCACCCACAGGAGAGCAAAATGGACAAAGTGGC	540
481	CCAAATCCGGAGAGGCTATGATTTATCCACCCACAGGAGAGCAAAATGGACAAAGTGGC	540
541	ATCTGCATAACTTTCAAGTACCAACAGTATGAGGCAATGATTCCTCAGCTCTTGGTCTGC	600
541	ATCTGCATAACTTTCAAGTACCAACAGTATGAGGCAATGATTCCTCAGCTCTTGGTCTGC	600
601	AACCCCAAGCTATATAGCTGCTCCATCCAGGCCACCTTTCCACAGAGAGCTCATTCACATG	660
601	AACCCCAAGCTATATAGCTGCTCCATCCAGGCCACCTTTCCACAGAGAGCTCATTCACATG	660
661	CCCCAGCTGTGCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACACACTT	720
661	CCCCAGCTGTGCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACACACTT	720
721	CAGTCGGCCACAGGACAAAAATTCCTCCATTTTGCAAGTCGATTCCTTTCTTGACGAC	780
721	CAGTCGGCCACAGGACAAAAATTCCTCCATTTTGCAAGTCGATTCCTTTCTTGACGAC	780
781	ATCTTTGACGCTGGATGGCTCAACGGCTGAAGACACACTTGTGTAAACAGAAACCTGGCAG	840
781	ATCTTTGACGCTGGATGGCTCAACGGCTGAAGACACACTTGTGTAAACAGAAACCTGGCAG	840
841	CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTCAATATATAAA	900
841	CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTCAATATATAAA	900
901	GCAAATTAATTTATCAGACACTCTTATTTCAAGTCTTTATCAAGATCACGCCAAGTGTGT	960
901	GCAAATTAATTTATCAGACACTCTTATTTCAAGTCTTTATCAAGATCACGCCAAGTGTGT	960
961	ATTTCCAAAAAGGGCCACAAAAATCGCTGCACATGTAATGGAGACCTTAATTCGGAGTCCA	1020
961	ATTTCCAAAAAGGGCCACAAAAATCGCTGCACATGTAATGGAGACCTTAATTCGGAGTCCA	1020
1021	CACCAAGCCTTCAGAGTGGAGGATTCATTGTGACCCAGAGATTCGCAAAATTTACCAAGCA	1080
1021	CACCAAGCCTTCAGAGTGGAGGATTCATTGTGACCCAGAGATTCGCAAAATTTACCAAGCA	1080
1081	TTTCAAGGATTTAGTATTTATCTATGAAGCTGTAAAGTAAACTTGGTGAAGGACACAGGT	1140
1081	TTTCAAGGATTTAGTATTTATCTATGAAGCTGTAAAGTAAACTTGGTGAAGGACACAGGT	1140
1141	ACTATCAATGAAACCTTGACAAATGGGTCTTCTCCATTACACCTCTTTTATATTTTAAA	1200
1141	ACTATCAATGAAACCTTGACAAATGGGTCTTCTTCCATTACACCTCTTTTATATTTTAAA	1200
1201	GGCCTGTGAATATCTTATTAACCTGCATATCACAAAAATAAAATATTTCTCTCATGTTT	1260
1201	GGCCTGTGAATATCTTATTAACCTGCATATCACAAAAATAAAATATTTCTCTCATGTTT	1260
1261	ACCAATTTA 1268	
1261	ACCAATTTA 1268	

```

RESULT 2
US-09-784B-4
; Sequence 4, Application US/09807784B
; Patent NO. 6653118
; GENERAL INFORMATION: Sei-ichi
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
; FILE REFERENCE: 210792

```

Query Match	85.1%;	Score 1079.6;	DB 4;	Length 1086;
Best Local Similarity	99.6%;	Pred. No. 0;	Indels	Gaps 0;
Matches 1082;	Conservative 0;	Mismatches	4;	
QY	34	ATGAAAACAGAAATGATGGCAGACTGCTAAGAACATCCTTGTGCTTGTCTTCTCTTGGC	93	
Db	1	ATGAAACAGAAATGATGGCAGACTGCTAAGAACATCCTTGTGCTTGTCTTCTCTTGGC	60	
QY	94	CTCTTTGGGGTGTGGGGCAGCAACAATTTTCATGCGAATGAAAGGGAAGCTGTG	153	
Db	61	CTCTTTGGGGTGTGGGGCAGCAACAATTTTCATGCGAATGAAAGGGAAGCTGTG	120	
QY	154	GACTGGTTTACTTTTATAAGTTACCTTAAAGACAAAAACAAGGAAAGTGGAGAGACTGGG	213	
Db	121	GACTGGTTTACTTTTATAAGTTACCTTAAAGACAAAAACAAGGAAAGTGGAGAGACTGGG	180	
QY	214	TTAGAGTACCTGTACTAGACTCTACAACTAGAGCTGGAGGAGAGTGAACAATATG	273	
Db	181	TTAGAGTACCTGTACTAGACTCTACAACTAGAGCTGGAGGAGAGTGAACAATATG	240	
QY	274	AATGACACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATGAGCATATGCGCTCT	333	
Db	241	AATGACACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATGAGCATATGCGCTCT	300	
QY	334	AAGAGTAAACAACACAGCCCTACTTAATAATACAAATGATGAGTCCCTTAAACCTGTGAATTAC	393	
Db	301	AAGAGTAAACAACACAGCCCTACTTAATAATACAAATGATGAGTCCCTTAAACCTGTGAATTAC	360	
QY	394	AGTAGAAATATGGACACACCAAGGTTTACTGCTGTGGAAACAGAGTTTCAAGGGTTCGG	453	
Db	361	AGTAGAAATATGGACACACCAAGGTTTACTGCTGTGGAAACAGAGTTTCAAGGGTTCGG	420	
QY	454	CTGATTCTATCCATCCCTCAGTTTCCCTCAAATCCGGAAGAGGCTATGATTATCCACCC	513	
Db	421	CTGATTCTATCCATCCCTCAGTTTCCCTCAAATCCGGAAGAGGCTATGATTATCCACCC	480	
QY	514	ACAGGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACAGTATGAG	573	
Db	481	ACAGGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACAGTATGAG	540	
QY	574	GCAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	633	
Db	541	GCAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	600	
QY	634	ACCTTTTCAACGAGAGCTCAATTCACATGCCCCCAGCTGTGCAACAGGGCCAGCTCATCAGAG	693	
Db	601	ACCTTTTCAACGAGAGCTCAATTCACATGCCCCCAGCTGTGCAACAGGGCCAGCTCATCAGAG	660	
QY	694	ATTTCCTGGAGGCTCTCTACACACATTGAGTCGGCCCGAGGACAAAATTCCTCCATTTT	753	
Db	661	ATTTCCTGGAGGCTCTCTACACACATTGAGTCGGCCCGAGGACAAAATTCCTCCATTTT	720	
QY	754	GCAAAGTCGGATTCTTTTCTTGACGACATCTTTTGACGACCTGGATGGCTCAACGGCTGAAG	813	
Db	721	GCAAAGTCGGATTCTTTTCTTGAGGCACTTTTGACGACCTGGATGGCTCAACGGCTGAAG	780	
QY	814	ACACACTTGTATAAGAAACCTGGCAGCGAAAAGAACAGAGCTTCTTCAAACTGCTCC	873	
Db	781	ACACACTTGTATAAGAAACCTGGCAGCGAAAAGAACAGAGCTTCTTCAAACTGCTCC	840	

QY 874 CTTCTTACATGCTCAATATATAAAGCAATTAATATACGACACTCTTATTTTCAGT 933
DB 841 CTTCTTACATGCTCAATATATAAAGCAATTAATATACGACACTCTTATTTTCAGT 900
QY 934 TCTTATCAAGATCAAGCAAGTGGTGATTTTCCAAAAGGCGACCAAAAATTCGTGGACA 993
DB 901 TCTTATCAAGATCAAGCAAGTGGTGATTTTCCAAAAGGCGACCAAAAATTCGTGGACA 960
QY 994 TGTATTGGAGACCTTAATTCGAGTCCACACAGAGCTTCAGAGTGAGAGTTCATTGTT 1053
DB 961 TGTATTGGAGACCTTAATTCGAGTCCACACAGAGCTTCAGAGTGAGAGTTCATTGTT 1020
QY 1054 ACCCAGAAATTTGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTATATCAATGAAAGCTGT 1113
DB 1021 ACCCAGAAATTTGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTATATCAATGAAAGCTGT 1080
QY 1114 AAGTAA 1119
DB 1081 AAGTAA 1086
RESULT 3
US-09-807-784B-2
; Sequence 2, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohara, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (213)..(1274)
US-09-807-784B-2
Query Match 53.1%; Score 673.8; DB 4; Length 1652;
Best Local Similarity 75.1%; Pred. No. 1.2e-206;
Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3;
QY 37 AAACAGAAAATGATGCAAGACTGCTAAGAACATCCTTTGGCTTTGCTCTTCTTGGCCCTC 96
DB 204 AGATAGAAATGACACCAAGCTCTAAGAACAGTTCTTTTGTCTTTCTTTTGGCCCTC 263
QY 97 TTTGGGCTCTGGGGCAGCAACAATTTTCATGAGAAATCAAGAGGAAAGCTGTGGAC 156
DB 264 TCTGGGCTCTGGGGCAGCAACAATTTTCATGAGAAATCAAGAGGAAAGCTGTGGAC 323
QY 157 TGGTTTACTTTTATAGTTTACTTAAAGACAAAACAAGAAAGTGGAGAGCTGGTTA 216
DB 324 TGGTTTACTTTTATAGTTTACTTAAAGACAAAACAAGAAAGTGGAGAGCTGGTTA 383
QY 217 GAGTACTCTTACTGACTCTCACTAGAGCTTGAAGCTGGAGAGAGTCACTAATGAT 276
DB 384 CAGTACTCTTACTGACTCTCACTAGAGCTTGAAGCTGGAGAGAGTCACTAATGAT 443
QY 277 GACACCAAGAGTGTTTTGGGAAGGACATTAACACAGCTATATGAAGCATATGCTTAAG 336
DB 444 AGCACCAGGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACATATTTCCAG 503
QY 337 ACTAACACACAGCTTATCTATATATCAATGATGGAGTCCCTTAAACCTGTGATTAAGT 396
DB 504 AATGACA---CAGCCTATCTAATATACACGATGGTGTCCCTGGATCTGTGAATTACAGC 560

QY 397 AGAAGCTATGGACACACCAAGGTTTACTGCTGTGGAAACAGAGTTCAAGGGTCTGGCTG 456
DB 561 AGACAGTATGGACATGCAAAAGGTTCTGCTGTATGGAACAGAAACGAGGGGTTCTGGCTG 620
QY 457 ATTCAATTCATCCTCAGTTTCTCCTCAATTTCCGGAAGAGCTATGATTTCCACCCACA 516
DB 621 ATACACTCTGTTCCCAAGTTTCCCAAGTTT-----ATGGCTATGATGATCCCAACCTCG 674
QY 517 GGGAGAGCAAAATGGACAAAGTGGCATCTGCATCAATTTCAAGTACAAACAGTATGAGCA 576
DB 675 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTTGGATACACAGAGTTTGGAA 734
QY 577 ATAGATTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAGCTGCTCCATCCCAAGCACC 636
DB 735 ATAGATTCTCAGCTCTTGGTCTTACCAACCAACATCTACAGCTGCTTCAATTTCAAAGCACC 794
QY 637 TTTTCCAGAGGCTCAATTCACATGCCCCAGCTGTGCACCAAGGGCCAGCTCATCAGAGATT 696
DB 795 TTTTCACTGGAACTTATCTACATGCCCGGATGTGCGCACTCCAGTTCCTTAAAGATC 854
QY 697 CTGGCAGGCTCTCCACCACTTCAGTCCGCCCGCCAGGCAAAAATTTCTTCCATTTTGCA 756
DB 855 CTTGTCCGGTACCTCGTGAACCTGCACTCAGCCCGAGGCTCTAAACTTCGTCCATTTGCA 914
QY 757 AAGTGGATTTCTTTTCTTGAAGCATCTTTGAGCTTGTGAGCTTGAATGGCTCAAGGCTGAAGACA 816
DB 915 AAATCAAGTTTATATCTGATGATCATCTTTTACAGGATGGATAGCTCAAAAAGTTTGAAGACA 974
QY 817 CACTTGTAAACAGAAACCTTGGCAGCGCAAAAGCAAGAGCTTCTTTCACAACTGCTCCCTT 876
DB 975 CATTGTGTAGCACAACCTTGGCAGAAAGAAACAGAGCTTCTTCAAACTGTTCCCTG 1034
QY 877 CTTTACATGTCTACAAATATAAAGCAATTAATATACGACACTCTTATTTTCAAGTTCT 936
DB 1035 CTTTACATGTCTACAAATATAAAGCAATTAATATACGACACTCTTATTTTCAAGTTCT 1094
QY 937 TATCAGATCAGCCCAAGTGTGTATTTCCCAAAAGGGCACCAAAAATCGTGGAGATGT 996
DB 1095 CGCACAAGCATTTCCAAATGGTGTGTTCATAAAGGGCTCCGCAAAATCGTGGAGCTGC 1154
QY 997 ATTGGAGACTAAATCGGAGTCCACACCAAGCTTCCAGAGTGGAGGANTCATTTGTACC 1056
DB 1155 ATTGGAGACTAAATCGGAGTCCACACCAAGCTTCCAGAGTGGAGGANTCATTTGTACC 1214
QY 1057 CAGAATTGGCAAAATTTACCAAGCATTTCAAGGATTTAGTATTTATGAAAGCTGTAAAG 1116
DB 1215 AAGAATCACTACATTTTACAGGCAATTTTCAAAATTTATATCTCCGTTATGGGTTCTGTAAG 1274
QY 1117 TAAACTTGTGAAAGGACACAGGTACTATCATTTGAAAACCTTGAATGGGTTCTTCTTC 1176
DB 1275 TAAACTCGGTGAAAGGCGCACCCCTCTGTCTTGTGAAAACACTGGCACTGGAAACATCTCGC 1334
QY 1177 ATTACACC--TTCTTTATATTTTAAAGGCTCTGTAAT 1211
DB 1335 CTTGATCTGTTCTCCATTAATTTCAAGGCTCTGTAAT 1371

RESULT 4
US-09-574-942-1
; Sequence 1, Application US/09574942
; Patent No. 6358723
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/574,942
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

Fri Oct 15 10:01:05 2004

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1575 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 MOLECULE TYPE: DNA
 US-08-639-294-1

Query Match 6.5%; Score 82.4; DB 3; Length 1575;
 Best Local Similarity 47.4%; Pred. No. 5.2e-16;
 Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;

142 GCGAAGAGCTGTGACCTGGTTTACTTTTATPAAGTT---ACCTAAAGACAAAACAGGAA 198
 155 GGGCAGCCTGTAGACTGGTTTCGTGCTCAAGCTGCCAGCTCTTAGAGGGTCGGGGAG 214
 199 AGTGGAGAGACTGGGTTAGAGTACCTGTACTAGACTCTAACTAGAACTGAAGCTGGAGGAAG 258
 215 GCGGCGCAGAGAGGGCTGCAGTATCTGGACGAGAGCTCCGAGGCTGGCGGGAC 274
 259 AGTGAGCAACTAATGAATGACACCAAGAGTGTGTTTGGAAAGGACATTACACAGCTATAT 318
 275 GGCAGGGCAGCTCATCAACAGCCCGAGGGGGCCCTGGCCGGAAGCCCTGCAGCCGCTGAC 334
 319 GAAGCATATGCTCTTAAGAGTAAACACACAGCCCTATCTAATATACATGATGAGTCCCT 378
 335 CGG-----AGCAACACACAGCCAGCTCGCTTCTCTTCAATGACCAACCGCCT 385
 379 AAAC-----TGGAATTACAGTAGAAGTATGACACCAAGGTTTACTGCTGTGG 432
 386 CAACCCAGCAAGGCTCAGGACTCTTCCATGCGTGGGACACAGAGGGTGTCTGCTCCTT 445
 433 AACAGAGTTCAAGGTTCTGCTGATTCATTCATTCCTCAGTTTCTCCAAATTCGGAA 492
 446 GACCAAGATGGGGCTTCTGCTGGTCCACAGTGTACCTTAACCTTCCCTCCACCGGCTCC 505
 493 ---GAAGGCTATGATTAATCCACCCACAGGGAGAGAAATGGACAAAGTGGGATCTGCATA 549
 506 TCTGCTGCATACAGCTGGCCTCATAGCGCTGTACTACGGGACAGACCTGCTCTGTGTG 565
 550 ACTTTCAAGTACACAGTATGAGCAATAGATTCTCAGCTCTTGTGCAACCCCAAC 609
 566 TCTTTTCCCTTTCGCTCAGTTCTCGAAGATGGGACAGCAGCTGACCTACACCTACCCCTGG 625
 610 GTCTATAGCTGCTCCATCCAGCCACCTTTTACACAGGAGCTCAATTCACATGCCCGAGTG 669
 626 GTCTATTAACACAGCTGGAAGGATCTTTGCGCCAGGAATTCCTCCGACTTGGAGAAATGTG 685
 670 TGCACCCAGGGCCAGCTCATFAGAGATTCCTGGCAGGCTCTCACCACACTTCAGTGGGCC 729
 686 GTCAAGGGCCACACGTTAGCCAAAGAACCTTGGAAACAGCAGCATCACATCCACAG 745
 730 CAGGGACAAAATTCCTCCATTTTGCAGAGTCGGATTCTTTTCTGACGACATCTTTGCA 789
 746 GCGCGGGCTGTTTTCAGAGCTTTCAGAGTTTCGCAAGTTTCAGCAAAATTTGGAGATGACCTGTACTCC 805
 790 GCTGGATGCTCAACGGCTGAAGACACACTTTGTTAAACGAAACCTGGCAGCGGAAAAAGA 849
 806 GGCTGTTGGCAGCAGCCCTTGGTACCACTGACCTGAGTCCAGTTCTGGCAGCAAAATCTGA 865
 850 CAAGAGCTTCTTCAAACTGCTCCCTTCCATGCTTACATGCTTCAATATATAAGCAATATAA 909
 866 GGCACTCTGCTCTTAACCTGCTCGGATATCTGGCAGGTTCTGAATGTGAACCAAGTAGCT 925
 910 TTATACGACA-----CTCTTATTTTCAAGTCTTATCAAGATCACCGCAGTGTGTATT 963
 926 TTCCCTGGACACCGGCCCAAGCTTCAACAGCAGAGGACCACTTCCAAATGTTGGGTG 985
 964 TCCCAAAAGGGCACCAAAAATCGCTGACATGTATTGGAGACCTTAATTCGAGTCCACAC 1023
 986 TCCCAAAAGGGGCC-----CTGGACCTGCTGGTGCATGATGAAATCGGAACCCAGGGA 1036

RESULT 6

US-08-639-294-1
 Sequence 1, Application US/08639294
 Patent No. 6265195
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Baron, Will F.
 TITLE OF INVENTION: Human DNase II
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESS: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/639,294
 FILING DATE: 25-Apr-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Johnston, Sean A.
 REGISTRATION NUMBER: 35,910
 REFERENCE/DOCKET NUMBER: P1024
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-3562
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168


```

; TYPE: DNA
; ORGANISM: Bos sp.
US-09-147-915-2

Query Match          5.6%; Score 71; DB 3; Length 927;
Best Local Similarity 47.7%; Pred. No. 1.8e-12;
Matches 317; Conservative 0; Mismatches 330; Indels 18; Gaps 3;

QY 405 TGGACACACAAAGTTTACTGCTGGGACACAGATTCAAGGTTCTGGCTGATTCTTC 464
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 TGGGCACACAAAGGTTGCTGCTCTCGACCAAGAGGGGGCTTCTGGTTGATCCACAG 134
QY 465 CATCCTCTAGTTTCTCCCAATCCCGA---AGAAAGCTATGATTATCCACCCACAGGAG 521
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 CGTTCCAAATCTCCCTCCACAGTCCCTCTCTGCTGGTACAGTGGCCTCTCTGGTCCCA 194
QY 522 ACGAATGGACAAAGTGGCATCTGCAATACTTTCAAGTACAAACAGTATGAGCAATAGA 581
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 AAATATGGGACAGCCCTGATCTGTGTAUCTTTTCTCTCACCCAGTCTCTGGATATCAG 254
QY 582 TTCTAGCTCTTGGTCTGCAACCCCAAGTCTATAGTGTCTCATCCAGCCACCTTCA 641
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 CAACAGCTGACCTATVACCTATCCACTGGTATATGACCAAGCTGGAAGGGGACTTTGG 314
QY 642 CCAGGAGCTCAATTCACATGCCAGCTGTGCACAGGGCCAGCTCATCAGAGATTCTGG 701
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 CAGAAATTCCTACTGAGGAGGTAGTCAAGGGCCATCAGTTGCGCCAGGACGGTG 374
QY 702 CAGGCTCTCCACACACTTCAGTCGGGCCAGGACAAATTCCTCATTTTCAAGTC 761
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 GAACAGCAGTGTAACTCATACATAAGAAAGAGGCCACATTCAGAGCTTTGCCAAAT 434
QY 762 GGAATCTTTTGTGACAGATCTTTGAGCCTGATGGCTCAAGGCTCAAGGCTGAAGACACTT 821
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 TGGAAATTTTGGAGATGACCTGTACTCTGGCTGGCTGGCGAAGCCCTTGGCAGTACCCT 494
QY 822 GTTAAAGAACTGGCAGGAAAGAGAGAGCTTCTTCAAACTGTCTCCCTTCTCTTA 881
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 GCAGGTCGAATCTGGCAAGATCTTCTGTGTATCTTGGCTTCCACTGCTCTGGGGCCCA 554
QY 882 CCAATGCTCAATATAAAGCAATTAATATCACGACACTCT-----TATTTCAGTTC 935
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555 GCATGATTTTGGCTGACTCAGACAGCTTTCCCTGGCCAGCTGGGCCAGCTTCAATGC 614
QY 936 TTATCAAGATCAGCCCAAGTGGTGTATTTCCCAAGGGGACCAAAATCGCTGGCAATG 995
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
615 CACAGAAGACCAATTCCAAGTGGTGTATACCCCAAAAGGGCC-----CTGGGCGTG 665
QY 996 TATTGGAGACCTAAATCGGAGTCCACACCAAGCTTCCAGAAAGTGGAGATTCTATTGTAC 1055
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
666 TGTGGTGAATGAATTCGGAACCAAGAGAGAGGACACCGGGGTGGGGGCACTCTGTGTGC 725
QY 1056 CCAGA 1060
Db      |||||
726 CCAGA 730

```

```

RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

```

```

; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)836-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match          4.1%; Score 51.8; DB 1; Length 7218;
Best Local Similarity 7.9%; Pred. No. 1e-05;
Matches 35; Conservative 217; Mismatches 189; Indels 0; Gaps 0;

QY 100 GGGGCTGCTGGGCGACCAAAATTCATGCAGAAATGAAGAGGAAAGCTGTGGACTGG 159
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1423 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1364
QY 160 TTTTACTTTTATAAGTTTACCTAAAGACAAACAAAGAGTGGAGAGACTGGTTAGAG 219
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1363 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1304
QY 220 TACTGTACTAGACTCTACAAGCTGAGAGAGTGGAGCAACTAATGAATGAC 279
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1303 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1244
QY 280 ACCAAGAGTGTTTTGGGAGGACATTACACAGCTATATGAAGCATATCTCTAAGAGT 339
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1243 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1184
QY 340 AACACACAGCTTATCTAATATACATGATGGAGTCCCTAAACCTGTGAATTACAGTAGA 399
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1183 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1124
QY 400 AAGTATGGACACACAAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTGATT 459
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1123 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1064
QY 460 CATTCATCCCTCAGTTTCTCCAAATTCGGAAGAAGGCTATGATTATCCACCACAGGG 519
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1063 GCAAGTCCCTCGACCTGAGCCAGCTCGAATTATCTGTGAGGCTATGCAACAGA 1004
QY 520 AGACGAAATGGACAAAGTGGC 540
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1003 AGGAAATAGTTATAGTAGC 983

```

```

RESULT 10
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976

```


NAME/KEY: intron
LOCATION: (136262)...(137936)
NAME/KEY: exon
LOCATION: (137937)...(138053)
NAME/KEY: intron
LOCATION: (138054)...(138637)
NAME/KEY: exon
LOCATION: (138638)...(138766)
NAME/KEY: intron
LOCATION: (138767)...(138864)
NAME/KEY: exon
LOCATION: (138865)...(138940)
NAME/KEY: intron
LOCATION: (138941)...(139765)
NAME/KEY: exon
LOCATION: (139766)...(139860)
NAME/KEY: intron
LOCATION: (139861)...(142245)
NAME/KEY: exon
LOCATION: (142246)...(142445)
NAME/KEY: intron
LOCATION: (142446)...(143605)
NAME/KEY: exon
LOCATION: (143606)...(143738)
NAME/KEY: intron
LOCATION: (143739)...(145838)
NAME/KEY: exon
LOCATION: (145839)...(145931)
NAME/KEY: intron
LOCATION: (145932)...(147385)
NAME/KEY: exon
LOCATION: (147386)...(147544)
NAME/KEY: intron
LOCATION: (147545)...(153274)
NAME/KEY: exon
LOCATION: (153275)...(153321)
NAME/KEY: intron
LOCATION: (153322)...(155088)
NAME/KEY: exon
LOCATION: (155089)...(155231)
NAME/KEY: intron
LOCATION: (155232)...(156025)
NAME/KEY: exon
LOCATION: (156026)...(156151)
NAME/KEY: intron
LOCATION: (156152)...(156826)
NAME/KEY: exon
LOCATION: (156827)...(156928)
NAME/KEY: intron
LOCATION: (156929)...(163399)
NAME/KEY: exon
LOCATION: (163400)...(163586)
US-09-676-610B-24

Query Match 3.1%; Score 39; DB 4; Length 169998;
Best Local Similarity 49.3%; Pred. No. 1.1;
Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 1052 GTACCCAGAAATGGCAAAATTTACCAAGCATTTCAAGGATTAGTATTATATCTATGAAAGCT 1111
Db 56948 GAACACCGTCTGGTAATTAAGATGACAGTTAAGCATTTATCTTCACTACCATAAGG 56889
QY 1112 GTAAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACAATGGGTCTT 1171
Db 56888 AAAAATAATCTTGCTTTGGCCACTTAATTAATTTGCTTACTTAATTTGAAATGTCCT 56829
QY 1172 CTTCATTACACCTCTTTTATTTTAAAGCCCTGTAATATCTTATATACCTTGCATATC 1231
Db 56828 TTCACTTTAAGAGTCTGTCTATTATTAAGTGGTATGTTTAAAGTAAAGTATGATTC 56769
QY 1232 ACAAAATAAAACATATTTCTCTCATGT 1258
Db 56768 CAAATATTAGATATCTTTAATCAATG 56742

RESULT 13
US-09-877-177A-10/c
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match 3.1%; Score 39; DB 4; Length 197496;
Best Local Similarity 49.3%; Pred. No. 1.2;
Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 1052 GTACCCAGAAATGGCAAAATTTACCAAGCATTTCAAGGATTAGTATTATATCTATGAAAGCT 1111
Db 64948 GAACACCGTCTGGTAATTAAGATGACAGTTAAGCATTTATCTTCACTACCATAAGG 64889
QY 1112 GTAAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACAATGGGTCTT 1171
Db 64888 AAAAATAATCTTGCTTTGGCCACTTAATTAATTTGCTTACTTAATTTGAAATGTCCT 64829
QY 1172 CTTCATTACACCTCTTTTATTTTAAAGCCCTGTAATATCTTATATACCTTGCATATC 1231
Db 64828 TTCACTTTAAGAGTCTGTCTATTATTAAGTGGTATGTTTAAAGTAAAGTATGATTC 64769
QY 1232 ACAAAATAAAACATATTTCTCTCATGT 1258
Db 64768 CAAATATTAGATATCTTTAATCAATG 64742

RESULT 14
US-09-134-000C-3209/c
; Sequence 3209, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3209
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3209

Query Match 2.7%; Score 34.6; DB 4; Length 1509;
Best Local Similarity 53.3%; Pred. No. 1.3;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 1103 ATGAAGCTGTAGTAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACA 1162
Db 1110 ATTTAAACCTTGAATTAACCTCTGCTAAAGAACCGGATATTTTCTTTGATTTTATTC 1051
QY 1163 ATGGGTCTTCTCCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATCTTATAAC 1222

Db 1050 AAAGACATTGACATTTTCACATTTTGACATAGTAAAGTTTCGGCAATAAATGTGTAGC 991
 QY 1223 CTGCATATCACAAATA 1239
 Db 990 AAACAAATAATAAAAAA 974

Db 1378 GAAAAATGGAGACCTTCAAGCATTTGGATAATCCGAATATCAATGCATCCAATGGTC 1437
 QY 188 AAAACAAGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACTAGACTCTACAACACTAGAA 247
 Db 1438 CACCCAGGCCGAGGATGAGTATGTGAATGAGCCACTGTACTCAACACCTTTGGCAACA 1497
 QY 248 GCTGGAGGAAGAGTGAGCAACTAATGAATGACA 280
 Db 1498 CCTTGGGAAAAAGCTGAGTACCTGAAGAAACA 1530

Search completed: October 14, 2004, 05:27:54
 Job time : 112.854 secs

RESULT 15
 US-08-484-438-5
 ; Sequence 5, Application US/08484438
 ; Patent No. 5811098
 ; Patent No. 5811098 5780031
 ; GENERAL INFORMATION:
 ; APPLICANT: Plowman, Gregory D.
 ; APPLICANT: Culouscou, Jean-Michel
 ; APPLICANT: Shoyab, Mohammed
 ; APPLICANT: Siegall, Clay B.
 ; APPLICANT: Hellistr m, Ingegerd
 ; APPLICANT: Hellistr m, Karl E.
 ; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,438
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/323,442
 ; FILING DATE: 14-OCT-1994
 ; APPLICATION NUMBER: US 08/150,704
 ; FILING DATE: 10-NOV-1993
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/981,165
 ; FILING DATE: 24-NOV-1992
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 5624-230
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3321 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 156..1782
 ; US-08-484-438-5

Query Match 2.7%; Score 34.6; DB 1; Length 3321;
 Best Local Similarity 51.6%; Pred. No. 2.2;
 Matches 79; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 128 GCAGAAATGAGAGGAGAGCTGTGACCTGTTTACTTTTATAGTTACCTAAAGAC 187

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:31:21 : Search time 67.2152 Seconds
(without alignments)
1694.809 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLRYGVCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_PUB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_PUB.pap.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1927	100.0	354	US-09-949-434-2	Sequence 2, Appli
2	1927	100.0	354	US-10-240-709-2	Sequence 2, Appli
3	1927	100.0	354	US-10-790-589-2	Sequence 2, Appli
4	1927	100.0	354	US-10-670-863-1	Sequence 1, Appli
5	1293.5	67.1	357	US-09-949-434-4	Sequence 4, Appli
6	1293.5	67.1	357	US-10-240-709-4	Sequence 4, Appli
7	1293.5	67.1	357	US-10-790-589-4	Sequence 4, Appli
8	1286.5	66.8	361	US-10-670-863-3	Sequence 3, Appli
9	596.5	31.0	360	US-10-408-167A-2	Sequence 2, Appli
10	557.5	28.9	348	US-10-408-765A-1164	Sequence 1164, Ap
11	111.5	5.8	818	US-10-437-963-145908	Sequence 145908,
12	111.5	5.8	832	US-10-437-963-145910	Sequence 145910,
13	101.5	5.3	475	US-09-729-674-142	Sequence 142, App
14	101.5	5.3	475	US-09-817-774-26	Sequence 26, Appl
15	101.5	5.3	475	US-10-429-160-32	Sequence 32, Appl

16	97.5	5.1	1886	12	US-10-147-299A-6	Sequence 6, Appli
17	97.5	5.1	2499	16	US-10-706-424-4	Sequence 4, Appli
18	94	4.9	1385	9	US-09-738-363-2	Sequence 2, Appli
19	94	4.9	1385	15	US-10-633-023-2	Sequence 2, Appli
20	94	4.9	2165	10	US-09-923-070A-29	Sequence 29, Appl
21	94	4.9	2165	10	US-09-923-070A-30	Sequence 30, Appl
22	94	4.9	2165	10	US-09-161-122-29	Sequence 29, Appl
23	94	4.9	2165	10	US-09-161-122-30	Sequence 30, Appl
24	94	4.9	2165	16	US-10-628	Sequence 393, App
25	93.5	4.9	361	9	US-09-925-300-1693	Sequence 1693, Ap
26	89	4.6	267	13	US-10-000-954-3	Sequence 3, Appli
27	88.5	4.6	2771	9	US-09-808-602-82	Sequence 70, Appl
28	88.5	4.6	2771	10	US-09-800-198-70	Sequence 45975, A
29	87.5	4.5	780	12	US-10-282-122A-45975	Sequence 103, App
30	87.5	4.5	3868	16	US-10-461-194-103	Sequence 163162,
31	87	4.5	607	12	US-10-425-114-67461	Sequence 63553, A
32	87	4.5	778	16	US-10-437-963-163162	Sequence 55, Appl
33	87	4.5	900	12	US-10-282-122A-63553	Sequence 54, Appl
34	87	4.5	964	12	US-10-042-865-55	Sequence 54, Appl
35	87	4.5	1045	12	US-10-029-020-54	Sequence 54, Appl
36	87	4.5	1045	12	US-10-038-854-42	Sequence 42, Appl
37	87	4.5	2613	16	US-10-038-854-40	Sequence 40, Appl
38	87	4.5	2628	16	US-10-038-854-38	Sequence 38, Appl
39	87	4.5	2721	16	US-10-038-854-36	Sequence 36, Appl
40	87	4.5	2725	16	US-10-038-854-36	Sequence 55910, A
41	86	4.5	424	12	US-10-282-122A-55910	Sequence 2, Appli
42	86	4.5	604	16	US-10-648-152-2	Sequence 37, Appl
43	86	4.5	618	16	US-10-648-152-37	Sequence 29, Appl
44	86	4.5	624	16	US-10-648-152-29	Sequence 63864, A
45	86	4.5	1005	12	US-10-425-114-63864	

ALIGNMENTS

RESULT 1

US-09-949-434-2
; Sequence 2, Application US/09949434
; Patent No. US20020028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-949-434-2

Query Match	100.0%;	Score 1927;	DB 9;	Length 354;
Best Local Similarity	100.0%;	Pred. No. 7.6e-189;		
Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLFFALSGVLGTPETSCRNEYGEADVDFIFYKLPKRTSKASEAGLOYL	60	
Db	1	MTAKPLRTVLSLFFALSGVLGTPETSCRNEYGEADVDFIFYKLPKRTSKASEAGLOYL	60	
QY	61	YLDSTRQTNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVGVSQYNSRQYG	120	
Db	61	YLDSTRQTNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVGVSQYNSRQYG	120	
QY	121	HAKGLLWNRTOGFWLIHSVPKPPVHGVEYPTSGRRYQGTGICITFGYSQFEIDFOLL	180	
Db	121	HAKGLLWNRTOGFWLIHSVPKPPVHGVEYPTSGRRYQGTGICITFGYSQFEIDFOLL	180	
QY	181	VLQPNITYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHSAOGLNFVHFPAKSGSFYT	240	

Db 181 VLOPNISYSCFIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT 240
Qy 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Db 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354

RESULT 2

US-10-240-709-2
; Sequence 2, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-240-709-2

Query Match 100.0%; Score 1927; DB 12; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Db 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTIQLHLYDTHNSTNDTAYLIYNDGVPGSVNYSQY 120
Db 61 YLDSTROTWNKSLYLINSTRSALGRTIQLHLYDTHNSTNDTAYLIYNDGVPGSVNYSQY 120
Qy 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSOFEEIDFQLL 180
Db 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSOFEEIDFQLL 180
Qy 181 VLOPNISYSCFIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT 240
Db 181 VLOPNISYSCFIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT 240
Qy 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Db 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354

RESULT 3

US-10-589-2
; Sequence 2, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Trustees of Dartmouth College

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-589-2

Query Match 100.0%; Score 1927; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Db 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTIQLHLYDTHNSTNDTAYLIYNDGVPGSVNYSQY 120
Db 61 YLDSTROTWNKSLYLINSTRSALGRTIQLHLYDTHNSTNDTAYLIYNDGVPGSVNYSQY 120
Qy 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSOFEEIDFQLL 180
Db 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSOFEEIDFQLL 180
Qy 181 VLOPNISYSCFIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT 240
Db 181 VLOPNISYSCFIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT 240
Qy 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Db 241 DDIFTGWIQAOKLTHLLAQWKKQELPNSCSLPYHVYNIKSIQVTSKSYFSSRQDHSK 300
Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYLRYGFCCK 354

RESULT 4

US-10-670-863-1
; Sequence 1, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-670-863-1

Query Match 100.0%; Score 1927; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60

Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 120
Db 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 120
QY 121 HAKGLLVNRTQGFMLIHSPKPPVHGYEYPTSGRRYGTGICITFGYSQPEIDFQLL 180
Db 121 HAKGLLVNRTQGFMLIHSPKPPVHGYEYPTSGRRYGTGICITFGYSQPEIDFQLL 180
QY 181 VLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
Db 181 VLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
QY 241 DDIPTGWIQAUKLTHLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 300
Db 241 DDIPTGWIQAUKLTHLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 300
QY 301 WCVSINGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 WCVSINGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354

RESULT 5
US-09-949-434-4
; Sequence 4, Application US/09949434
; Patent No. US2002028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-434-4

Query Match 67.1%; Score 1293.5; DB 9; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

Query Match 67.1%; Score 1293.5; DB 9; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 6
US-10-240-709-4
; Sequence 4, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-709-4

Query Match 67.1%; Score 1293.5; DB 12; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

Query Match 67.1%; Score 1293.5; DB 12; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 7
US-10-790-589-4
; Sequence 4, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357

Query Match 67.1%; Score 1293.5; DB 9; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

Query Match 67.1%; Score 1293.5; DB 9; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60
Db 1 MWARLRTSFALLFLGLFGLVGAANTISCRNEGKAVDWFIFYKLPKRNKESGETGLEVL 60
QY 61 YLDSTQTNKSLYLINSTRSALGRTOHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTOQLYEAAYASKNNTAYLIYNDGVPKPVNYSRKY 120
QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 177
Db 121 GHKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEIDF 180
QY 178 QLLVLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240
QY 238 FTYDDIFTCWIAQKLTLLAOTWQKKQELPNSCLPYHVNKISIGVTSKYSRSSRQD 297
Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYHVNKISIGVTSKYSRSSYQD 300
QY 298 HSKWCVSIKGSANRWTICIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLYLYRGCK 354
Db 301 HAKWCISQKGTNRWTICIGDLNRSLSHQAALRGGGFICTQNWQIYQAFQGLVLYESCK 357

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match      67.1%; Score 1293.5; DB 16; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEBKAVDWFTFYKLPKRNKESGETGLEYL 60

QY 61 YLDSTRTWNKSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119
Db 61 YLDSTTRSWKSKQLMNDTKSVLGRTOQLQYEAASKNNNTAYLIYNDGVPKPNYSRKY 120

QY 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFYQFEEIDF 177
Db 121 GHTKGLLLNWRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKKNQYEIDS 180

QY 178 QLVLPQNTYISCTIPSTFHWKLIYMPRMCAVSSSLKIPVRYLAELHSAQGLNFVHFKA 237
Db 181 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLLTTLQSAQOQKFLHFAKSD 240

QY 238 FYTDDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHYVNIKSGVTSKSYFSSROD 297
Db 241 SFLLDDIPAANWAQKLTHTLTETWQRKQELPNSCLPYHYVNIKALJSHRSYFSSYQD 300

QY 298 HSKWCVSIGKSANRWTICIGDLNLSLHQAALRGGGFICTKNHYIYQAFHKLRLYRGFCK 354
Db 301 HAKWCISQKTKNRTWTCIGDLNRSHPQAFRSGGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 8
US-10-670-863-3
; Sequence 3, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuka, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 3
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-670-863-3

Query Match      66.8%; Score 1286.5; DB 16; Length 361;
Best Local Similarity 66.1%; Pred. No. 4.4e-123;
Matches 236; Conservative 45; Mismatches 73; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 5 MMARLLRTSFALLFLGLFGVLGAATISCRNEBKAVDWFTFYKLPKRNKESGETGLEYL 64

QY 61 YLDSTRTWNKSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119
Db 65 YLDSTTRSWKSKQLMNDTKSVLGRTOQLQYEAASKNNNTAYLIYNDGVPKPNYSRKY 124

QY 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFYQFEEIDF 177
Db 125 GHTKGLLLNWRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKKNQYEIDS 184

QY 178 QLVLPQNTYISCTIPSTFHWKLIYMPRMCAVSSSLKIPVRYLAELHSAQGLNFVHFKA 237
Db 178 QLVLPQNTYISCTIPSTFHWKLIYMPRMCAVSSSLKIPVRYLAELHSAQGLNFVHFKA 237
```

```
Db 185 QLLVCPNVYSCSIPATFHQELIHPOLCTRASSEIPGRLLTTLQSAQOQKFLHFAKSD 244
QY 238 FYTDDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHYVNIKSGVTSKSYFSSROD 297
Db 245 SFLLDDIPAANWAQKLTHTLTETWQRKQELPNSCLPYHYVNIKALJSHRSYFSSYQD 304
QY 298 HSKWCVSIGKSANRWTICIGDLNLSLHQAALRGGGFICTKNHYIYQAFHKLRLYRGFCK 354
Db 305 HAKWCISQKTKNRTWTCIGDLNRSHPQAFRSGGGFICTQNWQIYQAFQGLVLYESCK 361

RESULT 9
US-10-408-167A-2
; Sequence 2, Application US/10408167A
; Publication No. US20030219428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/408,167A
; FILING DATE: 04-Apr-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/861034
; FILING DATE: 18-May-2001
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, David W.
; REGISTRATION NUMBER: NONE
; REFERENCE/DOCKET NUMBER: P1024D1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1739
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-408-167A-2

Query Match      31.0%; Score 596.5; DB 15; Length 360;
Best Local Similarity 38.0%; Pred. No. 3e-52;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGEAVDWFIFYKLPK-RTSKASEAGLQYLYLDSTRTWNKSL 73
Db 6 LAALLCPAGALTTCYDGSQGVDFVYVYKLPALRGSGEAAQGRGQYKLYLDESSGGRDGR 65

QY 74 YLINSTSALGRTOHLIYDTHNS--TNDTAYLIYND--GVPGSVNYSGRYGHAKGLIYVNR 131
Db 66 ALINSPGAVGRSLQPLY--RSNTSQLAFLLIYNDQPPQSKAQDSSWRGHTKGVLLIHDH 123

QY 132 QGFWLIHSPVKPP--VHGVEYPTSGRRYQGTGICITFYQFEEIDFOLLVLPQNIYS 188
Db 124 GGFVLIHSPVNFPPPPASSAAVSWPHSACTYQTLVCVSPFPFPAQFSKMGKQLTYTPWYN 183

QY 189 CFIPSTPHWKLIYMPRMCAVSSSLKIPVRYLAELHSAQGLNFVHFKA 248
```

Db 184 YOLEGFAOEFPPDENLVKHHVQSEWNSSTITTSQAGAVFQSFQKFKGDDLYSGWL 243
 QY 249 AOKLTHLLAQTWKKQKQELPNSCSLPYHYNISIGVTSKS--YFSSRODHKWCVSIIK 306
 Db 244 AALGTLNLOVQFVHKTVGILLPNSCSDIWOVLNVNQAIFPGFAGFSFNSTEDHSHKWCVPK 303
 QY 307 GSNRWTICIDNLSLHQAIRGGGFCITCKNHYIYOAFHKLILYRGFC 353
 Db 304 GP---WTCVDMNRNQEGEGGTICAQLPALWKAFOPLVKNYQPC 347

RESULT 10

US-10-408-765A-1164
 ; Sequence 1164, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1164
 ; LENGTH: 348
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-408-765A-1164

Query Match 28.9%; Score 557.5; DB 16; Length 348;
 Best Local Similarity 37.0%; Pred. No. 2.9e-48;
 Matches 129; Conservative 52; Mismatches 137; Indels 31; Gaps 9;

QY 17 LSGVLTGP--EISCRNEYGEAVDWFIFYKLPK-RTSKASEAGLQYLIDSTRQTNWKS 73
 Db 6 LAALCVPAAGALTCYGDSDQPDVFWVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65
 QY 74 YLINTSRALGRTLOHLYDTHNSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131
 Db 66 ALINSPGAVGSLQFLY--RSNTSOLAFLYNDQPPQSKAQDSMRGHTKGVLILDHD 123
 QY 132 QGFWLIHVPKFP--VHGVEYPTSGRRYGTGIC--ITFGYSQFEEIDFOLLVLPNI 186
 Db 124 GGFVLVHVSFVNPFPASSAAYSNPHSACTYGTGLCKQLTYTY-----FW 169
 QY 187 YSCFTPTSTHFWKLIYMRNCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYTDIFGT 246
 Db 170 YNYLEGIFAQFPPDENLVKHHVQSEWNSSTITTSQAGAVFQSFQKFKGDDLYSG 229
 QY 247 WIAOKLTHLLAQTWKKQKQELPNSCSLPYHYNISIGVTSKS--YFSSRODHKWCVS 304
 Db 230 WIAAALGNLQVQFVHKTVGILLPNSCSDIWOVLNVNQAIFPGFAGFSFNSTEDHSHKWCVS 289
 QY 305 IYGSANRWTICIDNLSLHQAIRGGGFCITCKNHYIYOAFHKLILYRGFC 353
 Db 290 PKGP---WTCVDMNRNQEGEGGTICAQLPALWKAFOPLVKNYQPC 335

RESULT 11

US-10-437-963-145910
 ; Sequence 145908, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 145908
 ; LENGTH: 818
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_46584C.1.1.pap
 ; US-10-437-963-145908

Query Match 5.8%; Score 111.5; DB 16; Length 818;
 Best Local Similarity 20.8%; Pred. No. 0.06;
 Matches 59; Conservative 33; Mismatches 111; Indels 81; Gaps 12;

QY 23 TPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYLIDSTRQTNWKSILYNSTRGA 82
 Db 250 TGFSCSGDPSSDLQVFWH-----GTRPYRSIVLDS---VW-----VSGKA 289
 QY 83 LGRTLOHLYDTHNSTNDTAYLIY--NDGVPGSVNSYR---QYGHAKGLLVNRTQGFWLI 137
 Db 290 YGSSISFPMYQTVNTQDEFYIYVTSQSP-----YRIMLDYTGTFRLLSWNNSSWAI 345
 QY 138 HVPKFPVHGYEYPTSGRRYGTGICITFGYSQFEEIDFOLLVLPNIYSCFIPSTFEW 197
 Db 346 YS-----QRPAAIGDCDPYSGCGPFGYCDFTSV-----IPRCQCPDGF- 384
 QY 198 KLIYMRNCANSSSLKIPVRYLAELHSAQGLNF-----VHFAKSSFYTDIFGTWIAQKL 252
 Db 385 -----PNSNSSSGCRRKQQLRCGEGNHFTWPMGKLPDKPFYVQD----- 425
 QY 253 KTHLLAQTWKKQKQELPNSCSLPYHVI-NKISIGVTSKSYFSSR 295
 Db 426 -----RSFECAECRNCSCATAYTNLTITGSPGTASQSR 463

RESULT 12

US-10-437-963-145910
 ; Sequence 145910, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 145910
 ; LENGTH: 832
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_46586C.1.1.pap
 ; US-10-437-963-145910

Query Match 5.8%; Score 111.5; DB 16; Length 832;

Db 349 YIFRVANHQKDLFRRTDGRCLWGRKPKVTECSYTSADGQRHSHKLLVSGFWGVARHFN 408
QY 315 IGDNLRSLSHQALR-GGGFICTKNHYIQAF-----HKLYLRYG 351
Db 409 VGDLMGSLAYCLACGGGHLDPFYIIYMAILLTHRCLRDDEHRCASKYG 456

RESULT 15

US-10-429-160-32
; Sequence 32, Application US/10429160
; Publication No. US20040023276A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Teresa R
; APPLICANT: Mao, Mao
; APPLICANT: Linsley, Peter S
; APPLICANT: Lund, Lund
; TITLE OF INVENTION: LXR Ligand Induced Genes and Proteins
; FILE REFERENCE: RS0200
; CURRENT APPLICATION NUMBER: US/10/429,160
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/377,714
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-429-160-32

Query Match 5.3%; Score 101.5; DB 16; Length 475;
Best Local Similarity 21.4%; Pred. No. 0.29;
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;
QY 6 LRTVLSLLFFA----LSGVLTGPEISCRNEYGEAVDF-----IFYKLPRKTSKASE 53
Db 40 LASVIFLLFAPFIVYFTFIMACDQXSALT-GPVVDIVTGCHARLSDIWAKTPPIITRKAQ 98
QY 54 EAGL-----QVLYLDSTRQWNKSL-----YLINSTRSALGRTLOH 89
Db 99 LYLTVTFQVLLYTSLPDFCHFLPGVYGGIPEGAVTPAGVNVKYQINGLQAWLLTHLEW 158
QY 90 LYDTHNSTNDTAYLYNDGVPGSVNSROYGHAKGLLVNRTQGFGLIHSVPKFPFVHGY 149
Db 159 FANAHLLSWFSPTIIFDNWIP-----LLWCAN---ILGYAVSTFAMVKGY 200
QY 150 EYPTSGRRYQGTGICITFGYSOFERIDFOLLVLQNIYSCF-----IPSTFHWKLIY 201
Db 201 FFFTSGARDCKFTG---NFFYNTMMGIEF-----NPRIGKWFDFLFFNGRPGIVAWTLIN 252
QY 202 MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHFAKSSPYTDDIFGTWIAQKLKTHLLAQ 259
Db 253 L-----SFAAK-----QRELSHVTNAMVLNVLQ-ALYVIDFF--W-----NE 288
QY 260 TWQKKKQEL-----PSNCSLP-----Y 276
Db 289 TWYLTIDICHDFGWLWGDCVWLPYLYTLQGLYLVVHPVQLSTPHAVGVLLGLVGY 348
QY 277 HVYNI-----KSGVTSKSYFSS--RQDHSKWCVS--IKGSANRWTC 314
Db 349 YIFRVANHQKDLFRRTDGRCLWGRKPKVTECSYTSADGQRHSHKLLVSGFWGVARHFN 408
QY 315 IGDNLRSLSHQALR-GGGFICTKNHYIQAF-----HKLYLRYG 351
Db 409 VGDLMGSLAYCLACGGGHLDPFYIIYMAILLTHRCLRDDEHRCASKYG 456

Search completed: October 14, 2004, 00:45:57
Job time : 68.2152 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:08:06 ; Search time 28.8776 seconds
(without alignments)
1179.175 Million cell updates/sec

Title: US-10-790-589-2
Perfect score: 1927
Sequence: 1 MTAQPLRTVLSLFFALS...KNHYIQAFHKLRLRYGFC 354
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: piri.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1373	71.3	356	2 JC7131	deoxyribonuclease
2	604.5	31.4	364	2 JE0205	deoxyribonuclease
3	596.5	31.0	360	2 JE0206	deoxyribonuclease
4	557.5	28.9	348	2 T45071	hypothetical prote
5	535	27.8	375	2 T19038	hypothetical prote
6	383	19.9	516	2 S40996	hypothetical prote
7	311	16.1	238	2 S44793	F0938.2 protein -
8	106	5.5	622	2 A36915	fructanase - Bacte
9	105	5.4	1445	1 A48148	protein-tyrosine-p
10	104	5.4	344	2 A44164	secreted glycoprot
11	97	5.0	668	2 T18635	hypothetical prote
12	94	4.9	1385	2 T18213	paraspinal crystal
13	94	4.9	2165	1 RNZA2	genome polyprotein
14	93.5	4.9	583	1 AJPWN2	asparagine synthas
15	91.5	4.7	774	2 AG1565	autolysin (amidase
16	91	4.7	1435	2 S54697	DNA polymerase III
17	90	4.7	1276	2 T18526	SREBP cleavage act
18	89.5	4.6	2825	2 T14271	Doc4 protein, stre
19	89	4.6	760	2 A99233	hypothetical prote
20	88.5	4.6	802	2 T21464	hypothetical prote
21	88.5	4.6	1309	2 F96509	hypothetical prote
22	88.5	4.6	1435	2 C90596	protein F27F5.19 I
23	88	4.6	3388	1 GNWVDP	hypothetical prote
24	87.5	4.5	175	2 S73337	genome polyprotein
25	87.5	4.5	471	2 JC1403	hypothetical prote
26	87.5	4.5	912	2 T02892	glutamate-ammonia
27	87	4.5	900	2 C64232	hypothetical prote
28	87	4.5	2109	1 A46309	alanine-tRNA ligas
29	86.5	4.5	1198	2 S51434	genome polyprotein
30	86.5	4.5	1198	2 S51434	hypothetical prote

30 86.5 4.5 1386 2 S73401 MG064 homolog R02-
31 85 4.4 867 1 GNLJSA pol polyprotein -
32 85 4.4 1087 2 T30330 gelsolin-related p
33 85 4.4 1442 1 B48148 protein-tyrosine-p
34 84.5 4.4 282 2 F71806 hypothetical prote
35 84.5 4.4 541 2 S57658 probable thioedox
36 84.5 4.4 566 2 T33042 hypothetical prote
37 84 4.4 675 2 S60612 protein-tyrosine k
38 84 4.4 5376 2 T42215 zonadhesin - mouse
39 83.5 4.3 394 2 JS0600 t-plasminogen acti
40 83.5 4.3 524 2 G64243 hypothetical prote
41 83.5 4.3 1064 1 S57450 protein-tyrosine k
42 83.5 4.3 1409 1 OFFPCP copia polyprotein
43 83.5 4.3 1436 2 T14895 DNA helicase 1 - A
44 83.5 4.3 1639 2 T50119 probable sensory t
45 83 4.3 415 2 B69875 conserved hypothet

ALIGNMENTS

RESULT 1

JC7131
deoxyribonuclease I (EC 3.1.21.1) - rat (strain Wistar)
N:Alternate names: deoxyribonuclease II-like acid deoxyribonuclease
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Dec-2002
C:Accession: JC7131
R;Tanuma, S.; Shiohara, D.
Biochem. Biophys. Res. Commun. 265, 395-399, 1999
A;Title: Cloning of a cDNA encoding a rat DNase II-like acid DNase.
A;Reference number: JC7131; MUID:20025354; PMID:10558878
A;Accession: JC7131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-356 <TAN>
A;Cross-references: GB:AF178974; NID:G6470130; PID:G6470131
C;Superfamily: deoxyribonuclease II
C;Keywords: hydrolase

Query Match 71.3%; Score 1373; DB 2; Length 356;
Best Local Similarity 71.0%; Pred. No. 1.1e-109;
Matches 252; Conservative 34; Mismatches 67; Indels 2; Gaps 1;
QY 1 MTAQPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDFYFKLPKRTSKASBEAGLQYL 60
Db 1 MTAQPLKAALPLLFVALSGVLGTFVISCINEDGKAVDFYFKLPKRTSGGGTGMGLDYL 60
QY 61 YLDSTQVWTKSLYLINSTRSALGRILQHLVDHNSNTDNTAVLIYNDVPGSVNYGRQYG 120
Db 61 YLDSTMTWTKSHHLLINSSRLGRTLEQLYEAHNAKNDTAVLIYNDVAPASVNYSGNYG 120
QY 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGOTGICITFGYQOFERIDFQ 178
Db 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGOTGICITFGYQOFERIDFQ 180
QY 179 LLVLPQNIYSCFIPSTFPHKLIYMPKNCANSSSLKIPRYLAELHSAQGLNFVHFAKSSF 238
Db 181 LLVLPQNIYSCFIPSTFPHKLIYMPKNCANSSSLKIPRYLAELHSAQGLNFVHFAKSSF 240
QY 239 YDDIFTGWIQAOKLTHLLAQTWQKKQBLPNCSLPHVYVNIKSTGVTSKSYFSRQDH 298
Db 241 YDDIFTGWIQAOKLTHLLAQTWQKKQBLPNCSLPHVYVNIKSTGVTSKSYFSRQDH 300
QY 299 SKWCYSIKGSANRWCIGDLNRSIHOALRGGGFICTKHYIYQAFHKLRLRYGFC 353
Db 301 SKWCYSTKDSQARWCIGDLNRSIHOALRGGGFICTKHYIYQAFHKLRLRYGFC 355

RESULT 2

JE0205
deoxyribonuclease II (EC 3.1.22.1) - pig
C:Species: Sus scrofa domestica (domestic pig)

C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Dec-2002
 C/Accession: JE0205
 R/Shiokawa, D.; Tanuma, S.
 Biochem. Biophys. Res. Commun. 247, 864-869, 1998
 A/Title: Cloning of cDNAs encoding porcine and human DNase II.
 A/Reference number: JE0205; MUID:98321218; PMID:9647784
 A/Accession: JE0205
 A/Molecule type: mRNA
 A/Residues: 1-364 <SHI>
 A/Cross-references: GB:AF060221; NID:g3309153; PIDN:AAAC39263.1; PID:g3309153
 C/Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides
 C/Superfamily: deoxyribonuclease II
 C/Keywords: hydrolase

Query Match 31.4%; Score 604.5; DB 2; Length 364;
 Best Local Similarity 37.9%; Pred. No. 5.9e-44;
 Matches 136; Conservative 60; Mismatches 144; Indels 19; Gaps 9;

QY 6 LRTVLSLTFALSGV-LGTPEISCRNEYGEAVDFWFIYKLPKRTSKAS-BEAGLQYLVD 63
 DB 1 MATLSPLLLAALLWVPGT--LTCYDGGQGVDFWVYKLPKRTSKAS-BEAGLQYLVD 58
 QY 64 STROTWNKSLYLINSTSALGRITLQHLVDTHNSTNDTAYLIYNDGVP---GSVNYSRQY 120
 DB 59 BESGWRDAGSINSITGALGRSLPLY---RNTSLAFLYNDQPPKYRGS-QHSSNRG 114
 QY 121 HAKGLAVNRTQGFWLIHSPKPPVH---GYEPTSGRRYGTGICITFGYSQPEIDF 177
 DB 115 HTKGVLDDGGFWLIHSPVNFPPSSAAAYSPWPPARTYGTGICITFGYSQPEIDF 174
 QY 178 QLLVLQPNYISCFPTSTPHWKLIIYMPKANSNLSKIPVRYLAELHSAQGLNFVHFVAKSS 237
 DB 175 QLTTPYPMYDYKLEGDFARFPYILEEVKGVHVLQEPWNSVILTSGAGASFOFACG 234
 QY 238 FYTDIDFTGTLAQKLTLLAQTQWKKQKQELPNSCLPIYHVNITKSGVTSKS---YFSSR 295
 DB 235 NFGDDLISGWLALAGSLNQLQFQWRSAGILPNSCSGVQHVLDVTQIAFPAGPAGNFAT 294
 QY 296 QDHKSKWCVSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIYQAFHKLILRYGFC 354
 DB 295 EDHSGKWCVA---PERPMTCVGMNRNKEEHRGGGTLCALPALWKAFOPLVKYQPC 350

RESULT 3
 JE0206
 deoxyribonuclease II (EC 3.1.22.1) - human
 C/Species: Homo sapiens (man)
 C/Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Dec-2002
 C/Accession: JE0206
 R/Shiokawa, D.; Tanuma, S.
 Biochem. Biophys. Res. Commun. 247, 864-869, 1998
 A/Title: Cloning of cDNAs encoding porcine and human DNase II.
 A/Reference number: JE0205; MUID:98321218; PMID:9647784
 A/Accession: JE0206
 A/Molecule type: mRNA
 A/Residues: 1-360 <SHI>
 A/Cross-references: GB:AF060222; NID:g3309154; PIDN:AAAC39852.1; PID:g3309155
 C/Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides
 C/Superfamily: deoxyribonuclease II
 C/Keywords: hydrolase

Query Match 31.0%; Score 596.5; DB 2; Length 360;
 Best Local Similarity 38.0%; Pred. No. 2.8e-43;
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--ETISCRNEYGEAVDFWFIYKLPK-RTSKASEEAGLQYLVDSTROTWNKSL 73
 DB 6 LAALICVPAGALTCTGDSGQVDFWVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65
 QY 74 YLINSTSALGRITLQHLVDTHNSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNWT 131
 DB 66 ALINSPGAVGRSLQPLY--RSNTSLAFLYNDQPPQPSKAQDSSMERGHTKGVLLDHD 123
 QY 132 QGFMLIHSVPKFP---VHGVEYPTSGRRYGTGIC--ITFGYSQFEIDFOLLVLQPN 186
 DB 124 GGFMLVHSPVNPFPDPASSAAYSWPHSACTYGTGLCKQLTYTY-----PWV 169
 QY 187 YSCFTPSFTHWKLIIYMPKANSNLSKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTG 246
 DB 170 YNYLEGIFAQFPPDENLVKGVHVSQEPWNSITLTQAGAVFQSFQKFKGDLISG 229
 QY 247 WIAQKLTHLLAQTQWKKQKQELPNSCLPIYHVNITKSGVTSKS---YFSSRQDHKWCVS 304
 DB 230 WLAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQTAFPPGAPGFSFNSTEDHSGKWCVS 289
 QY 305 IKGSRANRWTCIGDLNLSLHQAALRGGGFICTKHHYIYQAFHKLILRYGFC 353
 DB 290 PKGP---WTCVGMNRNKEEHRGGGTLCALPALWKAFOPLVKYQPC 335

RESULT 5
 T19038
 hypothetical protein C07B5.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

DB 66 ALINSPGAVGRSLQPLY--RSNTSLAFLYNDQPPQPSKAQDSSMERGHTKGVLLDHD 123
 QY 132 QGFMLIHSVPKFP---VHGVEYPTSGRRYGTGICITFGYSQFEIDFOLLVLQPN 186
 DB 124 GGFMLVHSPVNPFPDPASSAAYSWPHSACTYGTGLCKQLTYTYPWVYN 183
 QY 189 CFIPSTFHWKLIIYMPKANSNLSKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTGWI 248
 DB 184 YQLEGIFAQFPPDENLVKGVHVSQEPWNSITLTQAGAVFQSFQKFKGDLISGWL 243
 QY 249 AQKLKTHLLAQTQWKKQKQELPNSCLPIYHVNITKSGVTSKS---YFSSRQDHKWCVS 306
 DB 244 AAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQTAFPPGAPGFSFNSTEDHSGKWCVS 303
 QY 307 GSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIYQAFHKLILRYGFC 353
 DB 304 GP---WTCVGMNRNKEEHRGGGTLCALPALWKAFOPLVKYQPC 347

RESULT 4
 T45071
 hypothetical protein R31240.2 [imported] - human
 C/Species: Homo sapiens (man)
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Dec-2002
 C/Accession: T45071
 R/Lamerdin, J.; McCready, P.; Stillwagen, S.; Ramirez, M.; Carrano, A.
 submitted to the EMBL Data Library, November 1996
 A/Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region
 A/Reference number: 222906
 A/Accession: T45071
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-348 <LAM>
 A/Cross-references: EMBL:AD000092; PIDN:AA051172.1
 A/Experimental source: cell line 5HL2-B; fibroblast
 C/Genetics:
 A/Map position: 19p13.2
 A/Introns: 29/2; 89/3; 116/1; 159/1; 225/1
 C/Superfamily: deoxyribonuclease II

Query Match 28.9%; Score 557.5; DB 2; Length 348;
 Best Local Similarity 37.0%; Pred. No. 5.8e-40;
 Matches 129; Conservative 52; Mismatches 137; Indels 31; Gaps 9;

QY 17 LSGVLGTP--ETISCRNEYGEAVDFWFIYKLPK-RTSKASEEAGLQYLVDSTROTWNKSL 73
 DB 6 LAALICVPAGALTCTGDSGQVDFWVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65
 QY 74 YLINSTSALGRITLQHLVDTHNSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNWT 131
 DB 66 ALINSPGAVGRSLQPLY--RSNTSLAFLYNDQPPQPSKAQDSSMERGHTKGVLLDHD 123
 QY 132 QGFMLIHSVPKFP---VHGVEYPTSGRRYGTGIC--ITFGYSQFEIDFOLLVLQPN 186
 DB 124 GGFMLVHSPVNPFPDPASSAAYSWPHSACTYGTGLCKQLTYTY-----PWV 169
 QY 187 YSCFTPSFTHWKLIIYMPKANSNLSKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTG 246
 DB 170 YNYLEGIFAQFPPDENLVKGVHVSQEPWNSITLTQAGAVFQSFQKFKGDLISG 229
 QY 247 WIAQKLTHLLAQTQWKKQKQELPNSCLPIYHVNITKSGVTSKS---YFSSRQDHKWCVS 304
 DB 230 WLAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQTAFPPGAPGFSFNSTEDHSGKWCVS 289
 QY 305 IKGSRANRWTCIGDLNLSLHQAALRGGGFICTKHHYIYQAFHKLILRYGFC 353
 DB 290 PKGP---WTCVGMNRNKEEHRGGGTLCALPALWKAFOPLVKYQPC 335

RESULT 5
 T19038
 hypothetical protein C07B5.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

```

Qy 96 STNDTAYLLYNDGYPGSGVNTSRQYGHAKGLLVWARTQGFNLHSHVDPKPPVHGVEYPTSG 155
Db 280 NMNTTFSYMYNDEPDDSTIWSNSGSHAGKGVTVFQYTGFWMIHSTPKFSPKDMFRPSPNA 339
Qy 156 RRYGQTGICITFGYSGQFEEDIFQLLQLPNIYSFCIPSTFHWKLLIYMPRMCAANSSSLKIP 215
Db 340 HYYGQMGICISYNTVSLATTAQQLFYNTTFYQFNLQSFANQFPVLSQLKXKYEYKNSPP 399
Qy 216 VRYLAELHSAQGLMFVHFAKSSFYTDIDFTGWIAQKLKTHLLAQTWQKK--KQELPNSC 272
Db 400 LITSTKVLKSLGGQHFHFAXTGKGLYSDPVGFTLTKSSIKVETWNHQSDEYNLPSCV 459
Qy 273 SLPHVYVNIKISIGVTSKSY---PSSRODHSKWCYSIKGSANR----WTCIGDLNRL 322
Db 460 D-PNHVOSTWSAKYIRLPAYIDYSSYEDHSKFFVAYSESSKPPPIPVYICIGDINQV 515

RESULT 7
S44793
F09G3.2 protein - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C/Accession: S44793
R:Anderson, K.
submitted to the EMBL Data Library, February 1993
A/Description: Sequence of the C. elegans cosmid F09G8.
A/Reference number: S44776
A/Accession: S44793
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-238 <AND>
A/Cross-references: EMBL:L11247; NID:gl56280; PID:gl56283
C/Genetics:
A/Introns: 32/2; 171/3

Query Match 16.1%; Score 311; DB 2; Length 238;
Best Local Similarity 29.5%; Pred No. 4.3e-19;
Matches 76; Conservative 44; Mismatches 98; Indels 40; Gaps 10

Qy 12 LLPFALSGVLGTPEISCRNRYGEAVDWFIYKLPKRTSKASEAGLQVLYLDSTRQTWNK 71
Db 6 VLIFFSVITFTGNGKITQCKNMRKGSVDFVYVYKLPKLSAGTI--SGKERVYFDESSDWR 63
Qy 72 SYLYINSTRSALGRTLQHLIDYDTHNSTNDTAVLIYNDGVPGSVNTSRQY-GHAKGLLVNR 130
Db 64 G-NDINDPNVAVGATVSQVTSADKSN--FWFMYSDDDP--IKSADSYRGHAKGVSLFDS 118
Qy 131 TQGFNLHSPKPPVHGYEYPTSGRRYGGTGICITFGYSQFEIDFQLLVLPNIYSCF 190
Db 119 TTGFNLHSPNPPPKISFYSPNPAEKYGGSFCCASMEVQHLTEL----- 163
Qy 191 IPSTFHWKLLIYMPRMCAANSCLKIPRYVLAELHSAQGLNFVHFHAKSSFYTDIDFTGWIAQ 250
Db 164 ---AEHWKYIQ----ATPYIINPEKYATRFPTLKNVE---AKQSL-PRSAQTQFWISK 210
Qy 251 KLXT-----HLLAQTWQKK 264
Db 211 PIKTVOGVTLMAYAKHKK 228

RESULT 8
A36915
fructanase - Bacteroides fragilis
C/Species: Bacteroides fragilis
C/Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
C/Accession: A36915
R:Blatch, G.L.; Woods, D.R.
J. Bacteriol. 175, 3058-3066, 1993
A/Title: Molecular characterization of a fructanase produced by Bacteroides fra
A/Reference number: A36915; MUID:93259952; PMID:8491724
A/Contents: BP-1
A/Accession: A36915

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-622 <BLA>
A;Cross-references: GB:M83774; NID:g143370; PIDN:AAA22924.1; PID:g143372
A;Note: sequence extracted from NCBI backbone (NCBIN:1131947, NCBIP:1131949)
C;Superfamily: beta-fructofuranosidase

Query Match 5.5%; Score 106; DB 2; Length 622;
Best Local Similarity 21.8%; Pred. No. 0.5;
Matches 79; Conservative 34; Mismatches 142; Indels 108; Gaps 19;

QY 34 EAVDVFYKLPKRTKASEAGLOLYLDSTQRTWKSILYL---INSTRSALGRITLOHL 90
DB 84 EKVDYFVPALP---KGEKEVAVRIHLEKALCW-KELKLSDTFTTNDQYRPLYHH 138
QY 91 YDTHNSTNDYALILYNDGVPGSVNYRQYGHAKGLLVNNTQFGLIHSVPKPPVHYE 150
DB 139 TPLYGWNMDANGLVYKDGHYHLYFYQNPYFGSMWGNMHWG-----HSYSK-DLVH-WE 188
QY 151 Y-----PTSGRRYQGTGI---CITFGYSQFEIDFQLLVLPN-IYSCFIPSTFWKL 199
DB 189 HLEPALARDTLGHIFGSSVVDANTAGYGAGAIVAFYTSADKNGQIQCMAYSTNGRT 248
QY 200 IYMPRCANSSSLKIPVRYLAELHSAQGLNFVFAKSSFTYDDIFTGWI-----248
DB 249 F-----TKYEKNFV-----LTPFDGLKDFRDPKVFWAPD--QKWMVVSADKEMRF 293
QY 249 --AQKLTLLAATW-----OKKQELPNSCLPVLHYVNIKIGVTSKYSYFSSRODHSK 300
DB 294 YSSENLEKWTYMSGWGEYGVQSPQCEPDVPEL-----VDGNPDHKK 337
QY 301 WCVSISK-----GSANRWTCIGDLNLSLHQLRGGGTCT-----KHHYIY 340
DB 338 WALIVNVPCCYGGSATQY-FIGDFD-----GEKVCNDKPEVTKVGLDWGKHAT 388
QY 341 QAF 343
DB 389 VCF 391

RESULT 9
A48148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type gamma - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48148; C36065; S12051
R;Barnea, G.; Silvennoinen, O.; Shaanan, B.; Honegger, A.M.; Canoll, P.D.; D'Eustachio, Mol. Cell. Biol. 13, 1497-1506, 1993
A;Title: Identification of a carbonic anhydrase-like domain in the extracellular region
A;Reference number: A48148; MUID:93180796; PMID:8382771
A;Accession: A48148
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1445 <BAR>
R;Kaplan, R.; Morse, B.; Huebner, K.; Croce, C.; Howk, R.; Ravera, M.; Ricca, G.; Jaye, Proc. Natl. Acad. Sci. U.S.A. 87, 7000-7004, 1990
A;Title: Cloning of three human tyrosine phosphatases reveals a multigene family of receptors
A;Reference number: A36065; MUID:90384936; PMID:2169617
A;Accession: C36065
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 874-1118,1175-1409 <KAP>
R;Kreuger, N.X.; Streuli, M.; Saito, H. EMBO J. 9, 3241-3252, 1990
A;Title: Structural diversity and evolution of human receptor-like protein tyrosine phosphatases
A;Reference number: S12049; MUID:9106018; PMID:2170109
A;Accession: S12051
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 836-1406, M, 1408-1445 <KRU>
A;Cross-references: GB:X54132; NID:g35793; PIDN:CAA38067.1; PID:g35794
C;Genetics:

A;Gene: GDB:PTPRG; D3S1249
A;Cross-references: GDB:127351; OMIM:176886
A;Map position: 3p14.2-3p14.2
C;Superfamily: protein-tyrosine-phosphatase, receptor type gamma; carbonic anhydrase homone-phosphatase homology
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane protein-tyrosine-phosphatase homology <CAH>
F;60-321/Domain: carbonic anhydrase homology <CAH>
F;346-434/Domain: fibronectin type III repeat homology <3FN>
F;742-758/Domain: transmembrane #status predicted <TMN>
F;874-1108/Domain: protein-tyrosine-phosphatase homology <PTP>
F;1175-1399/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;1060/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1066/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.4%; Score 105; DB 1; Length 1445;
Best Local Similarity 18.4%; Pred. No. 1.8;
Matches 60; Conservative 51; Mismatches 89; Indels 126; Gaps 17;

QY 14 FFALSGVLG-----TPEISCRNEYGEAVDWFIYFKLPKRTSKASEAGLOLYLDSTQRT 68
DB 59 YWAYSGAYGPEHWVTSSVCGSRHQSPIDLDQY---ARVGEYQELQLDGFNDSNKT 115
QY 69 WNKSLYLINSTRLALQRTLOHLYDTNSTNDTAYLYNDGVPGSVNYSR---QYGHAKGL 125
DB 116 WKNV-----TGKTVAILLKDD-----YFVSGAGLPGRFKAKEVFEFHGHSNGS 158
QY 126 LVNWRQTGFWLHVSVPKFPVHGVEYPTSGRRYGGTGICITFGYSQFEEDFOLLVLPN 185
DB 159 A-----GSEHSINGRRF-----PVEMQIFPNPD 182
QY 186 IYSCFIPSTFWKL-----YMPRCANSSSLKIPVRYLAELHSAQGLNFVFAKSS 237
DB 183 DFDSPQTAISEWRIIGAMAIFFQVSPR---DNSALD-PI-----IHLKGV--VHEKET 231
QY 238 F-----YTDDIFTGWIQAOKLTHLAQTWQKKQELPNSCSPYH- 277
DB 232 FLDPFVRLDPLPASLGSYYRTGSLTTPCSEIVE-----WIVFRPVP-----ISYHQ 280
QY 278 ---VYNIKSIGVTSKYSYFSSRODHSK 300
DB 281 LEAFYSI-----FTTEQQDHVK 297

RESULT 10
A44164
secreted glycoprotein, 43 K - trichina
C;Species: Trichinella spiralis (trichina)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Sep-1997
C;Accession: A44164; A60630; S27862
R;Vassiliatis, D.K.; Despommier, D.; Misk, D.E.; Polvere, R.I.; Gold, A.M.; Van der Ploeg, J. Biol. Chem. 267, 18459-18465, 1992
A;Title: Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode Trichinella spiralis
A;Reference number: A44164; MUID:92406752; PMID:1382055
A;Accession: A44164
A;Molecule type: mRNA
A;Residues: 1-344 <VAS>
A;Cross-references: EMBL:M95499; NID:g162534; PID:g162535
A;Note: sequence extracted from NCBI backbone (NCBIN:113310, NCBIP:113321)
A;Note: part of this sequence was confirmed by protein sequencing
R;Gold, A.M.; Despommier, D.D.; Buck, S.W. Mol. Biochem. Parasitol. 41, 187-196, 1990
A;Title: Partial characterization of two antigens secreted by L1 larvae of Trichinella spiralis
A;Reference number: A60630; MUID:90377287; PMID:2398916
A;Accession: A60630
A;Molecule type: protein
A;Residues: 23, X, 25-62, YGSP, 66-80, T, 115-116, X, 118-132, 249-278 <GOL>
C;Keywords: glycoprotein

Query Match 5.4%; Score 104; DB 2; Length 344;
Best Local Similarity 19.0%; Pred. No. 0.35;
Matches 65; Conservative 60; Mismatches 119; Indels 98; Gaps 17;

QY 27 SCRNEYGEAVDWIFYK---LPKRTSKASEAGLOLYLDSTQRTWKSILYINSTRL 83

Db 26 TCKTATDD-TWELLKPKVGLLAKIISPANACWAN---DGANWNTDSGHALVQTLAEWM 81
QY 84 GRTLOHLYTHNSTNTAVLIYNDGPGSVNYSGRQGHAKGLLWNR--TQGFWLHISVP 141
Db 82 GPLT-----DDMTALGYSNTTPKS-TITSTSSKGLMFGNETTDGFWLLHTFE 130
QY 142 K-FPPVHGVEYPTSGRRYGTGICITFGYS-----QFEIDFQLLVQENIYSCFI 191
Db 131 RAFFNSVANSWSKFTSEGHMALCISEDNVPLIVPALQIQSV----- 174
QY 192 PSTFHWKLIYPMRC-----ANSSSL---KIPV-----RYLAELHSAQGLNFVHF 233
Db 175 -----VIYFGQVSSEKATEFADLTSLIDGSLPTTTPPLWNOQTITLNSA--LSTVVY 225
QY 234 AKSSFYTDIDFTGWIQAOKLTHLLAQTQWKKQELPSNCSLPHVNYIKSIGVTSKSYPS 293
Db 226 SKTSSRLMEYSGFLAKVMVNM--RIWAVTDNTLQTCGGKIGFVKVWSPVTDIGTON 283
QY 294 SR-QDHSKWCV-----SIKGSANRWTCI 315
Db 284 DRSKDKSQNAVIDDKPVFCFTNGYSTKQRTVAGSA---TCI 322

RESULT 11
T18635
hypothetical protein B0019.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18635
R:Kershaw, J.
submitted to the EMBL Data Library, November 1997
A:Reference number: Z19000
A:Accession: T18635
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-668 <WIL>
A:Cross-references: EMBL:AL008866; PIDN:CAAL5509.1; GSPDB:GN00019; CESP:B0019.1
A:Experimental source: clone B0019
C:Genetics:
A:Gene: CESP:B0019.1
A:Map position: 1
A:Introns: 42/3; 103/1; 144/2; 195/3; 258/1; 304/1; 357/1; 404/2; 457/2; 474/1; 589/1; 6

Query Match 5.0%; Score 97; DB 2; Length 668;
Best Local Similarity 20.0%; Pred. No. 3.2; Mismatches 98; Indels 136; Gaps 16;
Matches 69; Conservative 42; Mismatches 98; Indels 136; Gaps 16;
QY 5 PLRTVLSLLFFALSGVLGTPFISCRNEYGEAVDWFIFYKLPKRTSKASBEAGLQVLYLDS 64
Db 109 PTDLTQLLSLQELN-VRTTQQLTCGNN-----TVFQQRKKQL-----S 147
QY 65 FQTVNKSILY-LINSTRS-----ALGRFLQHLVD----- 92
Db 148 LQQQSTTLFTDLINSPELGLNTSVSAMSQMDTADSDSVNRMTQFFDAPGEQVPE 207
QY 93 -----THNSTNDTAVLIYNDGPGSVNYSGRQGHAKGLL-----VWNRQGFWLHIS 139
Db 208 IQALFTCSQNAFA-----VEILRRGHGOSLLAQQGMEVVRRLADGLLIYS 256
QY 140 VKPPPPVHGVEYPTSGRRYGTGICITFGYSQFEID-----FQLILVLA--ENIYSC 189
Db 257 QPAFAV---VQTSAGRRSGRQVIVAVPIPTLENIELVPAEAPFQQLIQNYGPTGHAY 313
QY 190 FTFPTFH---W-----KLIV-----NPRVCANSSSLK----- 213
Db 314 YFTMSFORATWRLNGRSGKVIYTSATGPLVWLTTDTTFAASCDSNSTSASLTWGLAHFS 373
QY 214 ---IPVRYLAELHSAQGLNFVHFAKS-----FYTDDIFTGWI 248
Db 374 YDVPETRRKLYQAIMVSLRFAFSPLDVSDVNFATDDLAKGTI 418

RESULT 12
T18213
parasporal crystal protein cry5Aal - Bacillus thuringiensis (fragment)
C:Species: Bacillus thuringiensis
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T18213
R:Narva, K.E.; Payne, J.M.; Schwab, G.E.; Hickie, L.A.; Galasan, T.; Sick, A.J.
submitted to the EMBL Data Library, July 1999
A:Description: Novel Bacillus thuringiensis microbes active against nematodes, and genes
A:Reference number: Z18829
A:Accession: T18213
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1385 <NAR>
A:Cross-references: EMBL:L07025; NID:g142868; PID:g142869; PIDN:AAA67694.1
C:Genetics:
A:Gene: CryVA (a)

Query Match 4.9%; Score 94; DB 2; Length 1385;
Best Local Similarity 20.5%; Pred. No. 15; Mismatches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;
QY 31 EYGEAVDWFIFYKLPKRTSKASEEAGLQVLYLDSR-----QTWNK----- 71
Db 399 EVQDSVETRLYQLP---AVDPQAGPNYVSIDSSNPILIQINMDTWKTPPQGASGWNWNL 454
QY 72 -----SLYLINSTRSALGR-----TLQHYDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVSGLSFLQDGTSLAGMGGGFADTIYSLPATH-----YLSYLYGTPTPYQTSNDY 507
QY 116 SRQYGHAKGLLVNRTQGFWLHISVPK---FPPVHGVEYPTSGRRYGTGICITFGYSQF 172
Db 508 SGHVGLVGV-----STPQEATLNIIG-----QPDQGNVSTMGF-PF 545
QY 173 BEIDFQLLVQPNIYSCFIPSTFHWKLIYMPKCANSSSLKIPVRYLALHS----- 224
Db 546 EKASYGGTVKKEWLANGANAMKLSPQSGIGIPITNTVTSGEYQIRCYASNDNTNVFNVDT 605
QY 225 -----AQGLNFVHFAKSSPYTDDIFTGWIQAOKLTHLLAQ----- 259
Db 606 GGANPIFQOINFASTVDNNTGVQANGVYVVKRSIATDTSFTETPAKTINVLHNTQSSD 665
QY 260 TWQKKQQLPNSCSLPHY-VYNIKISGVTSKSYFSSRQDHSKWCYSIKSANRWTCIGDL 318
Db 666 VFLDRIEFIPFLPLIYHGSYNTSS-GADDVLWSSNMNY--YDIIVNGQANS----SSI 718
QY 319 NRSLSHQALRG 328
Db 719 ASSMHLNKG 728

RESULT 13
RN2A2
genome polyprotein - human respiratory syncytial virus (strain A2)
N:Alternate names: polymerase L protein
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: human respiratory syncytial virus
A:Note: host Homo sapiens (man)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jun-2000
C:Accession: A40317; A28319; PS0048
R:Stec, D.S.; Hill III, M.G.; Collins, P.L.
Virology 183, 273-287, 1991
A:Title: Sequence analysis of the polymerase L gene of human respiratory syncytial virus
A:Reference number: A40317; MUID:9127488; PMID:2053282
A:Accession: A40317
A:Molecule type: mRNA
A:Residues: 1-2165 <STE>
A:Cross-references: GB:M75730; NID:g333955; PIDN:AAA47418.1; PID:g333956
R:Collins, P.L.; Olmsted, R.A.; Spriggs, M.K.; Johnson, P.R.; Buckler-White, A.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 5134-5138, 1987
A:Title: Gene overlap and site-specific attenuation of the viral polymerase
A:Reference number: A28319; MUID:87260943; PMID:2440043
A:Accession: A28319

QY	82	ALGR-----TIQHLYDTN-----STNDTAYLYNDGVPGSVNY	115
Db	87	QUSNHFRFGSCDVAHLYEEYGEDFVMDLGI FSEVPLDTRDNSYIVDAI	140
QY	116	SROYGHAKGLLWNRQTGGFWLHISVPK-----FPPVHGYYEYPTSG-RRYGQIGC	164
Db	141	----GVTSLYIGWGLDSVWISSEMGLNDDCBHFECFPFGCHLYSSKDSGFRW----	190
QY	165	ITGYQOFBEIDQLLVLPQNITYSCIPSTFH-----WKLHYPMRCAN-----	208
Db	191	-----YNPSTWSEAPYDPLALRHAFKAVKRLMTDVPFGVLLS	233
QY	209	---SSSL--KIPRYLAELHSAQ--GLNFVHFHAKSGFYTDIDFTG-WIAOKLKT--HLLA	258
Db	234	GLDSSLVASITSEYLAITYKAABONGSKLHSCVGLGSPDLKAGKEVADYLGTVHHEFT	293
QY	259	QTWQKKQBELPNSCSPYHV--YNIKSIGVTSKSYSSSRODHS---KWCVSIKGSANRW	313
Db	294	FTVQDGDIDATE--DVIIYHVETVDVTSIRASTPMFLMSRKIKSLGVKVVISGBGS----	345
QY	314	CIGDLNRSLHQAALRGGGGFCITCKNHYLYQAAPHK	345
Db	346	-----DEIFGG-----YLY--PHK	357
RESULT 15			
AG1565			
autolysin (amidase) homolog lin1064 [imported] - Listeria innocua (strain Clip1			
C/Species: Listeria innocua			
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001			
C/Accession: AG1565			
R/Glaser, P.; Burchgraben, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;			
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;			
J.; Jones, L.M.; Karst, U.			
Science 294, 849-852, 2001			
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournan,			
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.;			
A/Title: Comparative genomics of Listeria species.			
A/Reference number: AB1077; MUID:21537279; PMID:11679669			
A/Accession: AG1565			
A/Status: Preliminary			
A/Molecule type: DNA			
A/Residues: 1-774 <GUA>			
A/Cross-references: GB:AL592029; PIDN:CAC96295.1; PID:gl6413523; GSPDB:GN00178			
A/Experimental source: strain Clip1262			
C/Genetics:			
A/Gene: lin1064			
Query Match 4.7%; Score 91.5; DB 2; Length 774;			
Best Local Similarity 20.6%; Pred. No.12;			
Matches 68; Conservative 137; Indels 69; Gaps 18;			
QY	1	MTAKPLRTVLISLFFALSGVLGTPEISCRNEYGEAVDWFIFYKLPKRTSKA-SBEAGLQY	59
Db	363	LTITYPDSVVISKAVNLDDGITNPT-----GNGI-WTRAYKLEGTTSVAQAQYANQD	414
QY	60	LYLDSTQTNWKSILYLINSTRSALG----RTLOHYDTHNSTNDTAYLYNDGV--PGSVN	114
Db	415	VKISEVETQHTGYTINISIGGRAIGWLDKNAITLYDQAE-----YKNAVTLDDGTIK	465
QY	115	YSROYGHAKGLLWVN--RTQGFWLHISVPK-----PPVHGYYEYPTSGRRY	158
Db	466	-----NVQGNIAIWTEYFRTAGTKLVAQAEYANKVDQIVREAKTPRGTYYPKSGGNVI	519
QY	159	GQTGICITFGYSQFBEIDFOLLVLPQNITYSCIPSTHFWKLIYMP----RMCANSSSLK-I	214
Db	520	GLWLDV--KAPDMYDEITSKANVMEAVIE-NVEGNAVWTAPEYKSVGKLIIGPASPNDK	575
QY	215	PVRYLAELHSAQGLNFVHFHAKSGSFYTDIDFTGWLTAQK-LKTHLLAQTWOK-KKQELPNC	272
Db	576	TVKLTIRAQTSRGTY-----EFSSNGKTGVGLDKKAFKITYSTIYDESVVRDAVTNV	629

QY 273 S-----LPYHVNIKSIQVTSKSYFSSRQ 296
| | | | | : | | | | : : : |
Db 630 SGHTVWTLPYQYGVKSVG--SASNYNNKQ 657

Search completed: October 14, 2004, 00:32:21
Job time : 31.8776 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:11:29 ; Search time 71.1983 Seconds
(without alignments)
1404.834 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAFLRTVLSLFFALSGV.....KNHYIQAPKLYLRYGCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1927	100.0	354	4	AAB72416 Murine DN
2	1927	100.0	354	4	AAB72416 Murine DN
3	1293.5	67.1	357	4	AAB13014 Mouse deo
4	1293.5	67.1	357	5	AAB13015 Human deo
5	1286.5	66.8	361	4	ABG61821 Prostate
6	596.5	31.0	360	2	AAB72417 Human DNA
7	596.5	31.0	360	2	AAB72417 Human DNA
8	444	23.0	366	4	ABG63496 Drosophil
9	443	23.0	276	2	ABG63496 Drosophil
10	236.5	12.3	192	6	AAB72921 Bos tauru
11	101.5	5.3	475	2	ABG72082 Mouse end
12	101.5	5.3	475	2	AAY29333 Human sec
13	101.5	5.3	475	2	AAB93573 Human Del
14	101.5	5.3	475	4	AAB93572 Human Del
15	101.5	5.3	475	5	AAB39059 Human sec
16	101.5	5.3	475	6	ABU89736 Human pol
17	101	5.2	481	6	ABU89736 Protein d
18	97	5.0	419	2	ADA36462 Acinetoba
19	96	5.0	2165	5	AAB06703 Medium ch
20	95	4.9	2499	6	ABG67251 Respirato
21	94.5	4.9	2165	5	ABM70221 Phototrab
22	94	4.9	2165	5	ABG67249 Respirato
23	94	4.9	1385	2	AAR29516 Bt toxin
24	94	4.9	1385	2	AAR20066 B.thuring
25	94	4.9	1385	2	AAB59881 17a prote
					AAR29026 Bacillus

26	94	4.9	1385	2	AAR28803	Aar28803 Bt toxin
27	94	4.9	1385	2	AAR44201	Aar44201 Bacillus
28	94	4.9	1385	2	AAR58631	Aar58631 Bacillus
29	94	4.9	1385	2	AAR76112	Aar76112 PSI7a aca
30	94	4.9	1385	2	AAB13884	Aab13884 17a toxin
31	94	4.9	1385	3	AAB13891	Aab13891 Bacillus
32	94	4.9	1385	5	AAB10359	Aab10359 Bacillus
33	94	4.9	2165	2	AAR94930	Aar94930 RSV RNA-d
34	94	4.9	2165	2	AAR77015	Aar77015 Respirato
35	94	4.9	2165	4	AAB31881	Aab31881 Amino aci
36	94	4.9	2165	5	ABG67235	Abg67235 Respirato
37	94	4.9	2165	5	ABG67237	Abg67237 Respirato
38	94	4.9	2165	5	ABG67252	Abg67252 Respirato
39	94	4.9	2165	5	ABG67257	Abg67257 Respirato
40	94	4.9	2165	5	ABG67244	Abg67244 Respirato
41	94	4.9	2165	5	ABG67248	Abg67248 Respirato
42	94	4.9	2165	5	ABG67233	Abg67233 Respirato
43	94	4.9	2165	5	ABG67236	Abg67236 Respirato
44	94	4.9	2165	5	ABG67228	Abg67228 Respirato
45	94	4.9	2165	5	ABG67232	Abg67232 Respirato

ALIGNMENTS

RESULT 1

AAB72416

ID AAB72416 standard; protein; 354 AA.

XX AC AAB72416;

XX DT 03-MAY-2001 (first entry)

XX XX Murine DNase.

XX XX Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;

XX XX infectious disease.

XX OS Mus musculus.

XX PN WO200112793-A1.

XX PD 22-FEB-2001.

XX PF 01-MAY-2000; 2000WO-JP002893.

XX PR 17-AUG-1999; 95JP-00230870.

XX PA (TANU/) TANUMA S.

XX PI Tanuma S, Shiokawa D;

XX DR WPI; 2001-218348/22.

XX PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage

XX PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful

XX PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX PS Claim 3; Page 48-49; 61pp; Japanese.

XX CC The present sequence is a murine deoxyribonuclease (DLAD), which is an

XX CC endonuclease. DLAD is capable of divalent cation-independent cleavage of

XX CC DNA under acidic conditions. This protein can be used as a substitute for

XX CC DNase I in treating cystic fibrosis, and is useful in the prevention and

XX CC treatment of infectious diseases

SQ Sequence 354 AA;

Query Match 100.0%; Score 1927; DB 4; Length 354;

Best Local Similarity 100.0%; Pred. No. 9.1e-191;

Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPPEISCRNEYGEAVDWFIYKLPKRTSKASEAGLQYL 60
 Db 1 MTAKPLRTVLSLLFFALSGVLGTPPEISCRNEYGEAVDWFIYKLPKRTSKASEAGLQYL 60
 QY 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120
 Db 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120
 QY 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEEDFOLL 180
 Db 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEEDFOLL 180
 QY 181 VLOPNYISCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNPFVHAKSFYT 240
 Db 181 VLOPNYISCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNPFVHAKSFYT 240
 QY 241 DDIFTGWIQAOKLTHLLAQTWKKQELPNSCLPYHVNIKSGIVTSKSYFSSRDHSHK 300
 Db 241 DDIFTGWIQAOKLTHLLAQTWKKQELPNSCLPYHVNIKSGIVTSKSYFSSRDHSHK 300
 QY 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCK 354
 Db 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCK 354

RESULT 2
 AAE13014
 ID AAE13014 standard; protein; 354 AA.
 AC AAE13014;
 XX
 DT 28-JAN-2002 (first entry)
 DE Mouse deoxyribonuclease (DNase) II beta protein.
 XX
 KW Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
 KW cystic fibrosis; pulmonary.
 XX
 OS Mus sp.
 XX
 FN WO200175082-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US010635.
 XX
 PR 03-APR-2000; 2000US-00541840.
 PR 19-MAY-2000; 2000US-00574942.
 XX
 PA (DART-) DARTMOUTH COLLEGE.
 XX
 PI Eastman AR, Krieser RJ;
 XX
 DR WPI; 2001-662972/76.
 DR N-PSDB; AAD21288.
 XX
 PT New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
 PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.
 XX
 PS Claim 5; Page 14-15; 21pp; English.
 PS
 CC The invention relates to deoxyribonuclease (DNase) II beta proteins and
 CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA
 CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce
 CC their viscosity. The present sequence is mouse DNase II beta protein
 CC their viscosity. The present sequence is mouse DNase II beta protein
 XX
 SQ Sequence 354 AA;
 Query Match 100.0%; Score 1927; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 9.1e-191;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPPEISCRNEYGEAVDWFIYKLPKRTSKASEAGLQYL 60

Db 1 MTAKPLRTVLSLLFFALSGVLGTPPEISCRNEYGEAVDWFIYKLPKRTSKASEAGLQYL 60
 QY 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120
 Db 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLDYTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120
 QY 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEEDFOLL 180
 Db 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEEDFOLL 180
 QY 181 VLOPNYISCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNPFVHAKSFYT 240
 Db 181 VLOPNYISCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNPFVHAKSFYT 240
 QY 241 DDIFTGWIQAOKLTHLLAQTWKKQELPNSCLPYHVNIKSGIVTSKSYFSSRDHSHK 300
 Db 241 DDIFTGWIQAOKLTHLLAQTWKKQELPNSCLPYHVNIKSGIVTSKSYFSSRDHSHK 300
 QY 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCK 354
 Db 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCK 354

RESULT 3
 AAE13015
 ID AAE13015 standard; protein; 357 AA.
 AC AAE13015;
 XX
 DT 28-JAN-2002 (first entry)
 DE Human deoxyribonuclease (DNase) II beta protein.
 XX
 KW Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
 KW cystic fibrosis; pulmonary; chromosome 1p22.
 XX
 OS Homo sapiens.
 XX
 FN WO200175082-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US010635.
 XX
 PR 03-APR-2000; 2000US-00541840.
 PR 19-MAY-2000; 2000US-00574942.
 XX
 PA (DART-) DARTMOUTH COLLEGE.
 XX
 PI Eastman AR, Krieser RJ;
 XX
 DR WPI; 2001-662972/76.
 DR N-PSDB; AAD21289.
 XX
 PT New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
 PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.
 XX
 PS Claim 5; Page 16-17; 21pp; English.
 PS
 CC The invention relates to deoxyribonuclease (DNase) II beta proteins and
 CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA
 CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce
 CC their viscosity. The present sequence is human DNase II beta protein. The
 CC human DNase II beta gene is located at chromosome 1p22
 XX
 SQ Sequence 357 AA;
 Query Match 67.1%; Score 1293.5; DB 4; Length 357;
 Best Local Similarity 66.4%; Pred. No. 4.8e-125;
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPPEISCRNEYGEAVDWFIYKLPKRTSKASEAGLQYL 60

Db 1 MWARLRTSFGALLFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPKQKESGETGLEVL 60
 QY 61 YLDSTFQWTKNSLYLINSRSGALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSROY 119
 Db 61 YLDSTTRSRKSEQLMNDTKSVLGRTOQLYEAASKSNNTAYLIYNDGVPKPVNSRYK 120
 QY 120 GHAKGLLVNRTQGGFWLHISVPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 177
 Db 121 GHTKGLLLNWRVQGFLLHISIPQFPPIPEGVDYPTGRRNQSGGLCITFKYNQVEAIDS 180
 QY 178 QLLVLPNTYSCFIPSTFTHWKLIVMPKANSLSKIPVRYLAELHSAQGLNFVHFAKSS 237
 Db 181 QLLVCPNVYSCSIPATFHOELIHPOLCTRASSSEIPGRLTTTLOSAQGGKFLHFAKSD 240
 QY 238 FYTDDIFTGWAQKLKTHLLAQTKQKQLPNSCNSLPVHVYNIKSIQVTSKSYSSROD 297
 Db 241 SFLLDDIFAAMWAQRLKTHLLTETWQKQELPNSCNSLPVHVYNIKAIKLSRHSYSSYOD 300
 QY 298 HSKWCVSIGKSANRWTCIGDLNRSJHQAALRGCGFICTKNHYIYQAFHKLILRYGCK 354
 Db 301 HAKWCISQKGTNRWTCIGDLNRSJHQAALRGCGFICTQNWQIYQAFQGLVLYYESCK 357

RESULT 4

ABG61821
 ID ABG61821 standard; protein; 357 AA.

AC ABG61821;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated protein #22.

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

OS Mammalia.

PN WO200230268-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US032045.

PR 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.

PR 16-MAR-2001; 2001US-0276791P.

PR 06-APR-2001; 2001US-0276888P.

PR 24-APR-2001; 2001US-0281922P.

PR 30-APR-2001; 2001US-0286214P.

PR 04-MAY-2001; 2001US-0288589P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

DR N-PSDB; ABK92136.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,

XX useful for diagnosing prostate cancer (PC) or screening modulators of PC,

XX by determining if prostate cancer-associated genes are expressed in a

XX prostate tissue.

PS Claim 27; Page 317; 436pp; English.

XX The present invention relates to methods of detecting a prostate cancer-

XX associated transcript in a cell from a patient. The method comprises

XX contacting a biological sample from the patient with prostate cancer-

XX associated polynucleotides (designated PC genes) that selectively

CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins
 XX
 SQ Sequence 357 AA;

Query Match 67.1%; Score 1293.5; DB 5; Length 357;

Best Local Similarity 66.4%; Pred. No. 4.8e-125;

Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFPALSGVLGTPRISCRNEEGKAVDMFTFYKLPKQKESGETGLEVL 60

Db 1 MWARLRTSFGALLFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPKQKESGETGLEVL 60

QY 61 YLDSTQWTKNSLYLINSRSGALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSROY 119

Db 61 YLDSTTRSRKSEQLMNDTKSVLGRTOQLYEAASKSNNTAYLIYNDGVPKPVNSRYK 120

QY 120 GHAKGLLVNRTQGGFWLHISVPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 177

Db 121 GHTKGLLLNWRVQGFLLHISIPQFPPIPEGVDYPTGRRNQSGGLCITFKYNQVEAIDS 180

QY 178 QLLVLPNTYSCFIPSTFTHWKLIVMPKANSLSKIPVRYLAELHSAQGLNFVHFAKSS 237

Db 181 QLLVCPNVYSCSIPATFHOELIHPOLCTRASSSEIPGRLTTTLOSAQGGKFLHFAKSD 240

QY 238 FYTDDIFTGWAQKLKTHLLAQTKQKQLPNSCNSLPVHVYNIKSIQVTSKSYSSROD 297

Db 241 SFLLDDIFAAMWAQRLKTHLLTETWQKQELPNSCNSLPVHVYNIKAIKLSRHSYSSYOD 300

QY 298 HSKWCVSIGKSANRWTCIGDLNRSJHQAALRGCGFICTKNHYIYQAFHKLILRYGCK 354

Db 301 HAKWCISQKGTNRWTCIGDLNRSJHQAALRGCGFICTQNWQIYQAFQGLVLYYESCK 357

RESULT 5

ABG72417

ID AAB72417 standard; protein; 361 AA.

AC AAB72417;

DT 03-MAY-2001 (first entry)

DE Human DNase.

XX Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;

XX infectious disease.

XX Homo sapiens.

XX WO200112793-A1.

PD 22-FEB-2001.

XX 01-MAY-2000; 2000WO-JP002893.

XX 17-AUG-1999; 99JP-00230870.

XX (TANU/) TANUMA S.

XX Tanuma S, Shiokawa D;

XX WPI; 2001-218348/22.

XX N-PSDB; AAF60716.

AAW37920
ID AAW37920 standard; protein; 360 AA.
AC AAW37920;
XX
XX
DT 11-SEP-1998 (first entry)
DE Homo sapiens DNase II.
XX
XX DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
KW controlled cell death; apoptosis; metaplasia; cell turnover;
KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
KW autoimmune disease; diagnosis; anticancer agents;
KW enhanced chromosomal rearrangement; chromosome instability.
XX
XX Homo sapiens.
XX
XX WO9816659-A1.
PN
XX
PD 23-APR-1998.
XX
XX 09-OCT-1997; 97WO-US018262.
XX
XX 15-OCT-1996; 96US-0028539P.
PR
XX (DART-) DARTMOUTH COLLEGE.
PA
XX Eastman A, Krieser R;
PI
XX WPI; 1998-251301/22.
DR N-PSDB; AAV29137.
XX
XX Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
PT e.g. to induce apoptosis in tumour cells, identify agents modulating
PT apoptosis and digest DNA in sputum of cystic fibrosis patients.
XX
XX Claim 5; Page 18; 29pp; English.
XX
XX The sequence is that of deoxyribonuclease II (DNase II). The enzyme is
CC useful to digest DNA, e.g. in the lung sputum of cystic fibrosis patients
CC to reduce sputum viscosity. It (or fragments) are also useful to produce
CC antibodies, e.g. to study DNase II expression in cells. DNase II has
CC recently been linked with DNA fragmentation in the early stages of
CC controlled cell death (apoptosis), a process critical to homeostasis
CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead
CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
CC whilst too little can lead to cancer or autoimmune diseases. The
CC antibodies raised against it can therefore be used to diagnose apoptotic
CC stages in selected cells, by contacting cells with the antibody,
CC detecting binding of the antibody with DNase II and determining DNase II
CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic
CC agents, e.g. anticancer agents to promote apoptosis in cells. The
CC antibodies can also be used to identify cells susceptible to premature
CC death. The cDNA encoding it is useful to identify agents modulating
CC apoptosis in cells, by treating cells with an agent, transfecting cells
CC with cDNA and monitoring apoptosis compared with untreated cells.
CC Inhibitors identified may be useful in preventing diseases relating to
CC enhanced chromosomal re-arrangement. Vectors comprising the cDNA can be
CC used to induce apoptosis in selected cells, e.g. tumour cells or cells
CC involved in autoimmune disorders. Antisense oligonucleotides can be
CC administered to cells to inhibit DNase II expression to reduce chromosome
CC instability associated with cancer
XX
XX Sequence 360 AA;

Query Match 31.0%; Score 596.5; DB 2; Length 360;
Best Local Similarity 38.0%; Pred. No. 9.8e-53;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;
17 LSGVLGTP--EISCENEYGEAVDWFIFKLPK-RYSKASEAGLOLYLDSTRQWNKSL 73
6 LAALLCVFAGALTTCYGDGSGQPDWFFVVKLPALRGSGEAAQRLQYKYLDSSGGWRDGR 65

74 YLINSTRSALGRTLOHLYDTHNSTNDTAYLIYND--GVPSVNYSRQYGHAKGLVWNR 131
DB ALINSEPEGAVGRSLQPLY--RNTSOLAFLLYNDQPPQPSKQADSSMRGHTKGLVLLDHD 123
132 QGFVLIHSPKFP--VHGYYPTSGRRYGTGICITFGYSQFEIDFOLLVLPNIYS 188
DB GGFVLIHSPKFP--VHGYYPTSGRRYGTGICITFGYSQFEIDFOLLVLPNIYS 188
189 CRIPSTFWKLLYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYDDIFTGWI 248
DB YQLEGIFAQEPDLENVVKGHVSOEPMNSITLTSQAGAVFQSPAKFSKFGDLYSGWL 243
249 AQKLTALLAQWQKKQELPNSCSLPVHVYNIKSIQVTSKS--YFSSRQDHSKVCVSTK 306
DB AALGNTLNQVFWHKTIVGILPNSCSDIWOVLNVNQAIFPGPAGPSFSTEDHSHKVCVSPK 303
307 GSANRWTCIGDLNRLSHQALRGSGFICTKNHYIYQAFHKLRLRYGFC 353
DB WTCVGMNRNQGEEQGGTLCALPALWKAQFPLVKNYQPC 347
RESULT 8
ABB63496
ID ABB63496 standard; protein; 366 AA.
XX
AC ABB63496;
XX
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17280.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07599.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 17280; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 366 AA;

Query Match 23.0%; Score 444; DB 4; Length 366;
Best Local Similarity 31.0%; Pred. No. 6.6e-37;


```

XX PR 11-DEC-1997; 97US-0069205P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Aguilera RJ, Lyon CJ;
XX DR WPI; 2003-066229/06.
XX DR N-PSDB; ABS58030.
XX PT New nucleic acid and its encoded murine endonuclease-SR protein, useful
XX PT for treating tumors or cancers, e.g. leukemias or lymphomas, or for
XX PT designing and isolating peptidomimetics or inhibitors of Endo-SR for
XX PT treating these diseases.
XX PS Example 5; Col 82; 66pp; English.
XX CC The invention discloses an isolated nucleic acid molecule, which encodes
XX CC a murine endonuclease implicated in switch recombination (Endo-SR)
XX CC protein. The enzyme is useful at cleaving DNA at specific G-rich regions
XX CC which are implicated in modulating DNA rearrangements. Also disclosed are
XX CC methods for repairing DNA and modulating genetic recombination in a cell.
XX CC One example of genetic rearrangement is in mature B lymphocytes which can
XX CC alter their Ig constant region for another, significantly enhancing the
XX CC versatility of the immune system by allowing B cells to alter their
XX CC function without altering their ligand specificity. The endo-SR nucleic
XX CC acid and protein are useful (using gene therapy) for treating tumours or
XX CC cancers, e.g. leukaemias or lymphomas, as well as the modulation of
XX CC apoptosis and programmed cell death events. The endo-SR nucleic acid and
XX CC protein are also useful for designing and isolating peptidomimetics and
XX CC inhibitors of Endo-SR, which may be employed for treating the diseases.
XX CC The sequence presented is the mouse endo-SR protein
XX CC Sequence 192 AA;
XX SQ
Query Match 12.3%; Score 236.5; DB 6; Length 192;
Best Local Similarity 27.1%; Pred. No. 8.5e-16;
Matches 67; Conservative 27; Mismatches 96; Indels 57; Gaps 4;
QY 6 LRTVLSLFFALSGVLGTEPISCRNEYGEAVDWFIFVKLPKRTSKASEAGLQYLVDST 65
DB 1 MATLSLLLAALWV-PARALSCYDGGQFVDWVVKLPFAHSGSRDTPKGLTCYMQDN 59
QY 66 ROTWNSLYLINSRSLGRTTLQHLVDTHNSTNDTAYLYNDGVP--GSVNTSROYGHAK 123
DB 60 SDGWDQGVGYINSPGAVGSRSLQPLY--RKNSSQLAFLYNDQPKSSSTRSTGHTK 117
QY 124 GLLVNRRTQGFVLIHSPKFPVHGVEYPTSGRRYQGTGICITFGYSOFERIDFOLLV 183
DB 118 GK-----QLTVTY 125
QY 184 PNIYSCFIPSTFWKLIYMPRCANSSSLKIPRYLAELHSAQGLNFVHFPAKSSFYTDI 243
DB 126 PLVDYHKLKGFFAQKLPDLLETVIKQHLHPEWNSSVILTSQAGATFQSFARFGKGD 185
QY 244 FTGWIAQ 250
DB 186 YSGWLAE 192
RESULT 11
ID AAY29333
XX AC AAY29333 standard; protein; 475 AA.
XX AC AAY29333;
XX DT 29-SEP-1999 (first entry)
XX DE Human secreted protein clone pj323_2 protein sequence.
XX KW Human; secreted protein; nutrition; cytokine; cell proliferation;
XX KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
XX KW haematopoiesis regulation; tissue growth; inhibin; cadherin;

```

```

KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.
XX OS Homo sapiens.
XX PN WO9937674-A1.
XX PD 29-JUL-1999.
XX PF 21-JAN-1999; 99WO-US001404.
XX PR 22-JAN-1998; 98US-0072134P.
XX PR 20-JAN-1999; 99US-00235609.
XX PA (GEMY ) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Merberg D;
XX PI Treacy M, Agostino MJ, Steininger RJ, Wong GG, Clark HF, Fechtel K;
XX WPI; 1999-458682/38.
XX DR N-PSDB; AAX90448.
XX PT New polynucleotides encoding secreted human proteins derived from, e.g.
XX PT fetal brain potentially used as immunostimulators.
XX PS Claim 34; Page 126-128; 139pp; English.
XX CC The present sequence represents a human secreted protein. Human secreted
XX CC protein polynucleotides and proteins are predicted to have biological
XX CC activities which would make them suitable for treating, preventing or
XX CC ameliorating medical conditions in humans and animals, although no
XX CC supporting data is given. Suggested activities include nutritional
XX CC activity, cytokine and cell proliferation/differentiation activity,
XX CC immune stimulating (e.g. as vaccines) or suppressing activity,
XX CC haematopoiesis regulating activity, tissue growth activity,
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC activity, cadherin/tumour invasion suppressor activity, and tumour
XX CC inhibition activity. The polynucleotides are also stated to be useful for
XX CC gene therapy
XX SQ Sequence 475 AA;
Query Match 5.3%; Score 101.5; DB 2; Length 475;
Best Local Similarity 21.4%; Pred. No. 0.33;
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;
QY 6 LRTVLSLFFA---LSGVLGTEPISCRNEYGEAVDWF-----IFVKLPKRTSKASE 53
DB 40 LASVIFLLLEAPFIVYVYFIMACDQYSCALT-GPVVDIVTGHARLSDIWAKTPTTRKAQ 98
QY 54 EAGL---QYLYLDSTRTWNSL-----YLINSTRSALGRTLQH 89
DB 99 LYTLWVTQVLLYTSLPDFCHKFLPGYVGGIPEGAVTPAGVNVKQINGLQAWLITLLW 158
QY 90 LYDTHNSTNDTAYLYNDGVGVSQVNSROYGHAKGLVNRRTQGFVLIHSPKFPVHG 149
DB 159 FANAHLSSWFSPTIIFDNWIP-----LLWCAN---ILGYAVSTFAMVKGY 200
QY 150 EYPTSGRRYQGTGICITFGYSOFERIDFOLLVLOPNYSCE-----IPSTFWKLIY 201
DB 201 FFPTSDCKETG---NPFYVMGIEF-----NPRIGKWFDFKFFNGRPGIVAWTLIN 252
QY 202 MPRCANSSSLKIPRYLAELHSAQGLNFVHFPAKSSFYTDITFTGWIAQKLTLLAQ 259
DB 253 L-----SFAAK---QRELHSHVTNAMLVNLVQ-AIYVIDFF--W-----NE 288
QY 260 TWQKKQQL-----PSNCSLP-----Y 276
DB 289 TWYLTIDICHDFGWLGMGDCVWLPYLYTLOGLYLVTHFVQLSTPHAVGVLLGLVGY 348
QY 277 HVYNI-----KSIGVTSKYFSS--RQDHSKWCYS--IKGSNRWTC 314

```

Db	349	YIFRVANHQKDLFRRDGRCLWGRPKVIECSYTSADGQRHSHKLLVSGFWGVARHFN	408
QY	315	IGDLNLSLHQLR-GGGFTCKNHYYQAF-----HKLYLYRG	351
Db	409	VGDLMGSLAYCLACGGGHLFPYIIYMAILLTHRCLRDEHRCASKYG	456
RESULT 12			
AAW93573	ID	AAW93573 standard; protein; 475 AA.	
XX	AC	AAW93573;	
XX	DT	17-JUN-1999 (first entry)	
XX	DE	Human Delta7-sterol reductase protein.	
XX	KW	Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;	
XX	KW	diagnosis; screening; double bond removal; 7-dehydrocholesterol;	
XX	KW	organic polymeric ring; cholesterol.	
XX	OS	Homo sapiens.	
XX	PN	DE19739940-A1.	
XX	PD	18-MAR-1999.	
XX	PF	11-SEP-1997; 97DE-01039940.	
XX	PR	11-SEP-1997; 97DE-01039940.	
XX	PA	(GLOS/) GLOSSMANN H.	
XX	PI	Glosemann H, Moebius F, Fitzky B;	
XX	DR	WPI; 1999-191430/17.	
XX	DR	N-PSDB; AAX23387.	
XX	PT	Human Delta7-sterol reductase polypeptide - useful for diagnosis or	
XX	PT	treatment of genetic defects e.g. hereditary Smith-Lemli-Opitz syndrome.	
XX	PS	Disclosure; Page 29-30; 62pp; German.	
XX	CC	This invention describes a human Delta7-sterol reductase. The encoding	
XX	CC	DNA can be used to diagnose or correct human Delta7-sterol reductase gene	
XX	CC	defects e.g. hereditary Smith-Lemli-Opitz syndrome, and to produce the	
XX	CC	recombinant Delta7-sterol reductase polypeptide, which can be used to	
XX	CC	replace a defective Delta7-sterol reductase enzyme in humans or other	
XX	CC	animals. It is also useful to screen for Delta7-sterol reductase	
XX	CC	inhibitors or to introduce and remove double bonds in synthetic and	
XX	CC	naturally occurring organic polymeric ring systems (Delta 7-sterol	
XX	CC	reductase catalyses the conversion of 7-dehydrocholesterol to	
XX	CC	cholesterol)	
XX	QQ	Sequence 475 AA;	
Query Match			
Best Local Similarity 5.3%; Score 101.5; DB 2; Length 475;			
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;			
QY	6	LRTVLSLFFA---LSGLGTPTISCRNEYGEADVWF-----IFYKLPRKTSKASE	53
Db	40	LASVIFLLFAPFIVYVIMACDQYSCALT-GPVVDIVTGHARLSDIWAKTPTIRKAAQ	98
QY	54	EAGL-----QYLYLSTQTNKSL-----YLINSTRSALGRILQH	89
Db	99	LYTLWVTFQVLLYTSLPDFCHKFLPGYVGGIQEGAVTPAGVVKYQINGLQALLTHLLW	158
QY	90	LYDTHSNNDTAXLYLNDGVPSGVNYSQYGHAKGLLVNNTQGFVLIHSVPKPPVHG	149
Db	159	FANAHLLSWFPTIFDWNIP-----LLWCAN---ILGYAVSTAMVKG	200
QY	150	EYPTSGRRYGTGICITFGYSQFEIDFQLLVLPQNIYSCF-----IPSTFWKLIY	201

RESULT 15
 ABB55768
 ID ABB55768 standard; protein; 475 AA.
 AC
 AC ABB55768;
 XX
 XX 14-FEB-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 142.
 XX
 KW Human; clone b306-7; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
 XX
 OS Homo sapiens.
 XX
 XX US2001039335-A1.
 XX
 PD 08-NOV-2001.
 XX
 XX 04-DEC-2000; 2000US-00729674.
 XX
 PR 26-NOV-1997; 97US-0126425P.
 PR 04-DEC-1997; 97US-0067454P.
 PR 20-DEC-1997; 97US-0068379P.
 PR 02-JAN-1998; 98US-0070346P.
 PR 07-JAN-1998; 98US-0070643P.
 PR 08-JAN-1998; 98US-0070755P.
 PR 13-JAN-1998; 98US-0071304P.
 PR 22-JAN-1998; 98US-0072134P.
 PR 30-JAN-1998; 98US-0073095P.
 PR 18-FEB-1998; 98US-0075038P.
 PR 23-NOV-1998; 98US-00197886.
 PR 30-MAR-2000; 2000US-00539330.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE R R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEL/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX
 DR WPI; 2002-040725/05.
 DR N-PSDB; ABA90946.
 XX
 PT New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune disorders,
 PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or
 PT inflammations.
 XX
 PS Disclosure; Page 281-282; 349pp; English.
 XX
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (ABB55699-ABB55800), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and

CC proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
 CC clones b306-7 and yb8-1 respectively and the clones b306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytoprotective, anti-inflammatory, immunomodulator, vulnerary,
 CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease, myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus
 XX
 XX Sequence 475 AA;

Query Match 5.3%; Score 101.5; DB 5; Length 475;
 Best Local Similarity 21.4%; Pred. No. 0.33;
 Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;

QY 6 LRTVLSLLPFA---LSGVLGTPEISCRNEYGRAVDF-----IFYKLPRTSKASE 53
 DB 40 LASVIFLLFAPPVIVYFFIMACDQYSCALT-GEVVDIVTGHARLSDIWAKTPTITKAAQ 98
 QY 54 EAGL---QVLYLSDSTRQTWNKSL-----YLYNSTRSALGRTLQH 89
 DB 99 LYTLVTFQVLLYTSLPDFCHKFLPGYVGIGQAVTPAGVNVKYLINGLQALLHLWL 158
 QY 90 LYDTHNSTNDTAYLIYNDGVFGSVNYSRQYGHAKGLLVNRTQGFWLIIHSVPKPPVHG 149
 DB 159 FANAHLWSFPTIIPDNWIP-----LLWCAN---ILGYAVSTFAMVKGY 200
 QY 150 EYPTSGRRYQGTGCTTGFYSQFEEIDFQLLVLPNIYSCF-----IPSTFWKLIY 201
 DB 201 FFPTSARDCKFTG---NFFYNYMMGIEF---NPRIGKWFDFKLFNRPFGIVAWTLIN 252
 QY 202 MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHFASSEYTDITFTGWIQAQKLKTLAQ 259
 DB 253 L-----SFAK-----QRELHSHVTVNAMLVNVVLQ-AIVVIDFF--W-----NE 288
 QY 260 TWOKKKQEL-----PSNCSLP-----Y 276
 DB 289 TWYLTIDICHDFGWYLGWDCVWLPYLYTLQGLVLYVHPVOLSTPHAVGVLLGLVG 348
 QY 277 HVYNI-----KSGVTSKSYFSS--QDHSKWCVS--IKGSANEWTC 314
 DB 349 YIFRVANHQDLFRRTDGRCLIWGRKPKVTECSYTSADGQRHSHKLLVSGFWGVARHFN 408
 QY 315 IGLNLSLHQALR-GGGFICTKNHYIYQAF-----HKLYLRYG 351
 DB 409 VGDLMSLAYCLACGGGHLPLPYFIIMAILLTHRCLEDRHCASKYG 456

Search completed: October 14, 2004, 00:28:06
 Job time : 74.1983 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:24:00 ; Search time 31.3671 Seconds
(without alignments)
582.636 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLILRYGCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*
 - 2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*
 - 3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*
 - 4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:*
 - 5: /cgn2_6/ptodata/2/1aa/PTUS COMB.pep:*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1927	100.0	354	4	US-09-574-942-2
2	1927	100.0	354	4	US-09-807-784B-1
3	1293.5	67.1	357	4	US-09-574-942-4
4	1286.5	66.8	361	4	US-09-807-784B-3
5	596.5	31.0	360	3	US-09-147-915-3
6	596.5	31.0	360	3	US-08-639-294-2
7	596.5	31.0	360	4	US-09-861-034B-2
8	443	23.0	275	3	US-09-147-915-4
9	236.5	12.3	192	4	US-09-210-422-2
10	101	5.2	481	4	US-09-328-352-7749
11	94	4.9	1385	1	US-07-876-280-2
12	94	4.9	1385	1	US-07-675-772-2
13	94	4.9	1385	1	US-08-063-170-2
14	94	4.9	1385	1	US-08-158-232-2
15	94	4.9	1385	1	US-08-304-626-2
16	94	4.9	1385	1	US-08-316-301A-2
17	94	4.9	1385	2	US-08-611-928-2
18	94	4.9	1385	3	US-09-173-891-2
19	94	4.9	1385	3	US-09-076-137-2
20	94	4.9	1385	4	US-09-738-363-2
21	94	4.9	1385	5	PT-US92-03624-2
22	94	4.9	1385	6	5281530-1
23	94	4.9	1385	6	5426049-1
24	94	4.9	2165	1	US-08-514-975B-2
25	94	4.9	2165	5	PT-US95-12507-2
26	93.5	4.9	583	6	5256558-4
27	90	4.7	1445	1	US-08-015-986A-2

28	90	4.7	1445	2	US-08-446-363-2	Sequence 2, Appli
29	89	4.6	267	1	US-08-015-973-3	Sequence 3, Appli
30	89	4.6	267	1	US-08-448-164-3	Sequence 3, Appli
31	89	4.6	267	3	US-08-081-929-3	Sequence 3, Appli
32	86.5	4.5	473	4	US-09-284-768A-24	Sequence 24, Appli
33	86.5	4.5	1198	4	US-09-284-768A-10	Sequence 10, Appli
34	85	4.4	587	4	US-09-328-352-8096	Sequence 8096, Ap
35	85	4.4	1442	1	US-08-015-986A-3	Sequence 3, Appli
36	85	4.4	1442	2	US-08-446-363-3	Sequence 3, Appli
37	84	4.4	675	4	US-08-426-509A-4	Sequence 4, Appli
38	84	4.4	675	4	US-08-232-545-4	Sequence 4, Appli
39	84	4.4	675	5	PT-US95-05008-4	Sequence 4, Appli
40	83.5	4.3	335	4	US-09-134-001C-5021	Sequence 5021, Ap
41	82	4.3	393	4	US-09-107-532A-4397	Sequence 4397, Ap
42	82	4.3	520	3	US-08-810-003-5	Sequence 5, Appli
43	82	4.3	520	4	US-09-776-490-5	Sequence 5, Appli
44	82	4.3	1438	4	US-09-209-916-1	Sequence 1, Appli
45	82	4.3	1471	1	US-08-683-839B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-574-942-2

; Sequence 2, Application US/09574942

; Patent No. 6358723

; GENERAL INFORMATION:

; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/574,942

; CURRENT FILING DATE: 2000-05-19

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; TYPE: PRT

; ORGANISM: Mus sp.

US-09-574-942-2

Query Match	100.0%;	Score 1927;	DB 4;	Length 354;
Best Local Similarity	100.0%;	Pred. No. 7.6e-206;		
Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLFFALSGVLTPTPEISCRNBYGEADVDFIYKLPKRTSKASREAGLQYL	60	
Db	1	MTAKPLRTVLSLFFALSGVLTPTPEISCRNBYGEADVDFIYKLPKRTSKASREAGLQYL	60	
QY	61	YLDSTQTNKSLYLINSTRSALGRTLQHLXYDTHNSTNDTAYLIYNDGVPGSVNSRQYG	120	
Db	61	YLDSTQTNKSLYLINSTRSALGRTLQHLXYDTHNSTNDTAYLIYNDGVPGSVNSRQYG	120	
QY	121	HAKGLLVNRTQGFVLIHSVPKFPVHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL	180	
Db	121	HAKGLLVNRTQGFVLIHSVPKFPVHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL	180	
QY	181	VLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIIPVYLAELHSAQGLNFVHFAKSSFT	240	
Db	181	VLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIIPVYLAELHSAQGLNFVHFAKSSFT	240	
QY	241	DDIFTGWIAQKLKTHLLAQTWQKKQELFNSGSLPHYVNIKSIQVTSKSYFSSRQDHSK	300	
Db	241	DDIFTGWIAQKLKTHLLAQTWQKKQELFNSGSLPHYVNIKSIQVTSKSYFSSRQDHSK	300	
QY	301	WCVSIKGSANRWTCIGDLNRSIHOALRGCGGFTCKNHYIQAFHKLILRYGCK	354	
Db	301	WCVSIKGSANRWTCIGDLNRSIHOALRGCGGFTCKNHYIQAFHKLILRYGCK	354	

```

RESULT 2
US-09-807-784B-1
; Sequence 1, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shikawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-807-784B-1

Query Match      100.0%; Score 1927; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-206;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
;
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 120
Db 61 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 120
;
Qy 121 HAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDFQL 180
Db 121 HAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDFQL 180
;
Qy 181 VLQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
Db 181 VLQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
;
Qy 241 DDIFTGWIAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKSYFSSRODHSK 300
Db 241 DDIFTGWIAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKSYFSSRODHSK 300
;
Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
;

RESULT 3
US-09-574-942-4
; Sequence 4, Application US/09574942
; Patent No. 6358723
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/574,942
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-574-942-4

Query Match      67.1%; Score 1293.5; DB 4; Length 357;
Best Local Similarity 66.4%; Pred. No. 2.8e-135;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

```

```

Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
;
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 119
Db 61 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPKPNYSRKY 120
;
Qy 120 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 177
Db 120 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 177
;
Qy 121 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 180
Db 121 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 180
;
Qy 178 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 178 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
;
Qy 181 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
Db 181 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
;
Qy 238 FYTDDIFTGWIAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 297
Db 238 FYTDDIFTGWIAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 297
;
Qy 241 SFLDGIFAAWVAQRLKTHLLTETWQKRQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 300
Db 241 SFLDGIFAAWVAQRLKTHLLTETWQKRQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 300
;
Qy 298 HSKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 298 HSKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
;
Qy 301 HAKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 357
Db 301 HAKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 357
;

RESULT 4
US-09-807-784B-3
; Sequence 3, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shikawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 3
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-784B-3

Query Match      66.8%; Score 1286.5; DB 4; Length 361;
Best Local Similarity 66.1%; Pred. No. 1.7e-134;
Matches 236; Conservative 45; Mismatches 73; Indels 3; Gaps 2;

Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 5 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 64
;
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 119
Db 65 YLDSTROTWNKSLYLINSTRSALGRTLOHLYDTHNS-TNDTAYLIYNDGVPKPNYSRKY 124
;
Qy 120 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 177
Db 125 GHAKGLLVNRTQGWLIHSVPKPPV--HGVEYPTSGRRYGTGICITFGYSQFEIDF 184
;
Qy 178 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 185 QLLVQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 244
;
Qy 238 FYTDDIFTGWIAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 297
Db 245 SFLDGIFAAWVAQRLKTHLLTETWQKRQELPNSCLPYHYVNIKSGIVTSKSYFSSROD 304
;
Qy 298 HSKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 305 HAKWCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 361
;

```



```
; FILING DATE: 18-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: F1024D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3562
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-861-034B-2

Query Match          31.0%; Score 596.5; DB 4; Length 360;
Best Local Similarity 38.0%; Pred. No. 1.2e-57;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGEAVDFIFKLPK-RTSKASAEAGLQYLIDSTRQTWNKSL 73
DB 6 LAALLCPAGALICYGDSGGPWFVYKLPALRGSGEAAQRGLOKYLDDESSGWRDGR 65
QY 74 YLINSTSALGRYLOHLYDTHNSTNDTAYLIYND--GVPGSVNYSROYGHAKGLLVNRT 131
DB 66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQPSKAQDSSMRGHTKGVLLDHD 123
QY 132 QGFWLHVSVPKPP--VHGVEYPTSGRRYCGTGCITTCGYSOFEEIDFOLLVLQNIYS 188
DB 124 GGFWLHVSVEFPFPASSAAYSPHSACTYGGTLLCVSFPFAQFSKMGKQLTYTPWVYN 183
QY 189 CFTPTSPFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFHFAKSSFYTDIDFTGWI 248
DB 184 YQLEGFAQFPDLENVVGKHVYSQEPWNSITLTSQAGAVFQSPAKFSKFGDLDLSGWL 243
QY 249 AQKLTLLAQTQWKQKQELPNSCLSPYHYNIKSIGVTSKS--YFSSRODHKWCVSIIK 306
DB 244 AAALGTNLQVQFWHKTGVILPNSCSDIWLQVLNVNQLAFPGPAGPSPFNSTEDHSKWCVSPK 303
QY 307 GSNARWTCIGDLNRLSHQALRGGGFICTKKNHYIYQAEHKLILRYGFC 353
DB 304 GP---WTCVGDMMNRNOGEORGGGTLCALPALWKAFOPLVKNYQPC 347

RESULT 8
US-09-147-915-4
; Sequence 4, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krieser, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-147-915-4

Query Match          23.0%; Score 443; DB 3; Length 275;
Best Local Similarity 27.1%; Pred. No. 5.4e-18;
Matches 67; Conservative 27; Mismatches 96; Indels 57; Gaps 4;

QY 6 LRTVLILLFPALSGVLGTPDISCRNEYGEAVDFIFKLPKRTSKASAEAGLQYLIDST 65
DB 1 MATYRSILLALLLVV-PAEALSCVGDSPQVDFVYVVKLFAHSGSRDTPKGLTCKYWDQN 59
QY 66 RQTNKSLYLINSTRSALGRYLOHLYDTHNSTNDTAYLIYNDGVP--GSVNYSGROYGHAK 123
DB 60 SDGWQDGVGYNSPEGAVGRSLQPLY--RKNSSQLAFLYNDQPPKSSSTRDSTGHGHTK 117
QY 124 GLLVNRTQGFWLHVSVPKPPVHGVEYPTSGRRYCGTGCITTCGYSOFEEIDFOLLVLQ 183
DB 118 GK-----QLITYTY 125
QY 184 PNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFHFAKSSFYTDI 243
DB 126 PLVYDHKLEGGFAQKLPDLLETIVIKNQHLVHEPWNSSVILTSQAGATFQSPAKFGKFGDDL 185
QY 244 FTGWIQA 250
DB 186 YSGWLAE 192

RESULT 10
US-09-328-352-7749
; Sequence 7749, Application US/09328352
```

```
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7749
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7749

Query Match 5.2%; Score 101; DB 4; Length 481;
Best Local Similarity 23.2%; Pred. No. 0.029;
Matches 55; Conservative 27; Mismatches 83; Indels 72; Gaps 12;

QY 139 SVKFPFVHGVEYPTSGRRYGGTGCITFGYQFEEDFOLLVLPNIYSCPTPS-----193
Db 234 SQAKSPTEHG---SLAKDYAQOHPEVVKAYLQATIEANFLIQEPEKSEFIAEKGIP 290
QY 194 -----TFH-----WKLYMPRCANSSSLK-----IPVRYLAELH 223
Db 291 AEVVYLFHGPLGLQTRDLTWKPEYRKATQIAIDTLKVLGKNDGTLVDNKFIDDPQIKDAF 350
QY 224 SAQGLNF---VHFAPKSSFTYDDIFTCGTAQKLT-HLLAQTQKKQKQELPNSCSIPYHV 278
Db 351 QASGLNYSQALADYAKSPVANDALG---QPIKTFDQVTVIIVKGEKVRSS-YETPEHA 406
QY 279 YN-----IKSIGVTSKYSFSSRODHSK-----WCVS-----IKGSANRW 312
Db 407 FSDLKKIQAGKTVRVVYS--QDHQSDKILLANLAWATDKAGQIQAFLLKDDAEKW 461

RESULT 11
US-07-876-280-2
; Sequence 2, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 amino acids
```

```
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; INDIVIDUAL ISOLATE: PSI7
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
; US-07-876-280-2

Query Match 4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97;
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;

QY 31 EYGEADVDFIFKLPKRTSKASEAGLQVLYLDSTR-----QTWNK-----71
Db 399 EVQDSVETRLYGQLP-----AVDPQAGNIVSIDSSNPIIQINMDTWKTPPQAGSGWNTNL 454
QY 72 -----SLYLINSTRSALGR-----TLQHLDYDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVSGLSFQLQDGTFLSAGMGGGFADTIYSLPATH-----YLSYLYGTPYQTSINY 507
QY 116 SRQYGHAKGLLVNRTQGFWLHSVPK---FPPVHGVEYPTSGRRYGGTGCITFGYQSF 172
Db 508 SGHVGALVG-----STPQEATLENIIG-----QPDEQCNVSTMGF-PF 545
QY 173 EEDFDQLLYLPQNIYSCPTPSFTFWKLIYMPRCANSSSLKIPVRYLAELHS-----224
Db 546 EKASYGGTVVKEWLNGANAMKLSPGQSIGIPITNTVTSGEYQICRCRYASNDNTNVFNVD 605
QY 225 -----AQGLNFVHFAPKSSFTYDDIFTCGTAQKLT-HLLA-----259
Db 606 GGANPIFQIINEASTVDNNTGVQGGANGVIVVKSIIATDNTSFTPEIPAKTINVHLTNQSSD 665
QY 260 TWQKKQKQELPNSCSIPYH-VYNIKSGIVTSKYSFSSRODHSKQKVCISIKGSANRWTCIGDL 318
Db 666 VELDRIEFIPFSLPIYHGSYNTSS-GADDVLWSSSNMNY--YDIIVNGQANS-----SSI 718
QY 319 NRSLHQALRG 328
Db 719 ASSMILLNKG 728

RESULT 12
US-07-675-772-2
; Sequence 2, Application US/07675772
; Patent No. 5262399
; GENERAL INFORMATION:
; APPLICANT: Hickie, Leslie A.
; APPLICANT: Sick, August J.
; APPLICANT: Schwab, George E.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Payne, Jewel M.
; TITLE OF INVENTION: No. 5262399el Compositions and Methods for the Control of
; TITLE OF INVENTION: Flukes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROMAN SALIWANCHIK
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/675,772
```

FILING DATE: 19910327
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SALIWANCHIK, ROMAN
REGISTRATION NUMBER: 21,023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: BACILLUS THURINGIENSIS
INDIVIDUAL ISOLATE: PS17
IMMEDIATE SOURCE:
LIBRARY: LAMBDA GEM (TM) - 11 LIBRARY OF KENNETH NARVA
CLONE: 17A
US-07-675-772-2

Query Match 4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97; Indels 112; Gaps 16;
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;

Qy 31 EYGEAVDWFIFYKLPKRTSKASEAGLYLYLDSTR-----QTNK----- 71
Db 399 EVQDSVETRLYGQLP-----AVDPOAGPNVYSIDSSNPIIQINMDTWKTPPQAGSGWNTNL 454
Qy 72 -----SLYLINSTRSALGR-----TLQHLVDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVGLSFLQDGTRLSAGMGGFADTIYSLPATH-----YLSLYGTPYQTSNDY 507
Qy 116 SRQYGHAKGLLVNRTQGFWLHISVPK---PPPVHGYEYPTSGRRYGTGICITFGYSQF 172
Db 508 SGHVGLVGV-----STPQATLPLNIIG-----QPDEQGNVSTMGP-PF 545
Qy 173 BEIDFQLLVLPNIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHS----- 224
Db 546 EKASYGTVVKEWLGANAMKLSFGSIGIPITNTVTSGEYQIRCRYASNDNTNVFNVDT 605
Qy 225 -----AQGLNFVHFAKSSFYTDIDFTGWIQAOKLKTLLAQ----- 259
Db 606 GGANPIFQINFASTVDNNTGVGANGVYVVKSIATDTSFTEIPAKTINVHLTNQSSD 665
Qy 260 TWQKKQELPSNCSLPYH-VYNIKSIGVTSKSYFSSRQDHSKWCVSIKGSANRWTCIGDL 318
Db 666 VFLDRIFIPSLPLIYHGSYNTSS-GADDVLMSSNNY--YDLIVNGQANS-----SSI 718
Qy 319 NRSLHQALRG 328
Db 719 ASSMHLNKG 728

RESULT 13
US-08-063-170-2
Sequence 2, Application US/08063170
Patent No. 5350576
GENERAL INFORMATION:
APPLICANT: Kim, Leo
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Compositions and Methods for Inducing an Immune
TITLE OF INVENTION: Response for Protection Against Endoparasites and Exoparasites
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID R. SALIWANCHIK
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL

COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/063,170
FILING DATE: 19930517
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,141
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/759,248
FILING DATE: 13-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: SALIWANCHIK, DAVID R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/S 103.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: PS17
INDIVIDUAL ISOLATE: PS17a
US-08-063-170-2

Query Match 4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97; Indels 112; Gaps 16;
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;

Qy 31 EYGEAVDWFIFYKLPKRTSKASEAGLYLYLDSTR-----QTNK----- 71
Db 399 EVQDSVETRLYGQLP-----AVDPOAGPNVYSIDSSNPIIQINMDTWKTPPQAGSGWNTNL 454
Qy 72 -----SLYLINSTRSALGR-----TLQHLVDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVGLSFLQDGTRLSAGMGGFADTIYSLPATH-----YLSLYGTPYQTSNDY 507
Qy 116 SRQYGHAKGLLVNRTQGFWLHISVPK---PPPVHGYEYPTSGRRYGTGICITFGYSQF 172
Db 508 SGHVGLVGV-----STPQATLPLNIIG-----QPDEQGNVSTMGP-PF 545
Qy 173 BEIDFQLLVLPNIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHS----- 224
Db 546 EKASYGTVVKEWLGANAMKLSFGSIGIPITNTVTSGEYQIRCRYASNDNTNVFNVDT 605
Qy 225 -----AQGLNFVHFAKSSFYTDIDFTGWIQAOKLKTLLAQ----- 259
Db 606 GGANPIFQINFASTVDNNTGVGANGVYVVKSIATDTSFTEIPAKTINVHLTNQSSD 665
Qy 260 TWQKKQELPSNCSLPYH-VYNIKSIGVTSKSYFSSRQDHSKWCVSIKGSANRWTCIGDL 318
Db 666 VFLDRIFIPSLPLIYHGSYNTSS-GADDVLMSSNNY--YDLIVNGQANS-----SSI 718
Qy 319 NRSLHQALRG 328
Db 719 ASSMHLNKG 728

RESULT 14

```

US-08-158-232-2
; Sequence 2, Application US/08158232
; Patent No. 5596071
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Kennedy, M. Keith
; APPLICANT: Randall, John Brooks
; APPLICANT: Meier, Henry
; APPLICANT: Uick, Heidi Jane
; APPLICANT: Poncerrada, Luis
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Schwab, George E.
; APPLICANT: Fu, Jenny
; TITLE OF INVENTION: No. 5596071el Bacillus thuringiensis Toxins Active
; TITLE OF INVENTION: Against Hymenopteran Pests
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,232
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/887,980
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/797,645
; FILING DATE: 25-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/703,977
; FILING DATE: 22-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/SCJ104.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; STRAIN: PSI7
; INDIVIDUAL ISOLATE: PSI7a
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522 (pMYC1627) NRRL B-18651
US-08-158-232-2
Query Match 4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97;
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;
;
; 31 EYGRAVDWFIYKLPKRTSKASEAGLOLYLDSTR-----QTWNK----- 71
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 399 EQVDSVETLQQLP-----AVDPQAGPNVSLDSNPIQINMDTWKTPQAGSGWNTNL 454
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 72 -----SLYLINSTRSALGR-----TIQHLDYDTNSTNDTAYLIYNDGVP--GSVNY 115
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 455 MRGSVGLSFLQDGTLSAGMGCGFADTIYSLPATH-----YLSLYGTPTQTSQNY 507
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 SRQYGHAKGLLVNRTQGFWLIIHSPK---PPPVHGYEYPTSGRRYGTGICITFGYSQF 172
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 508 SCHVGALGV-----STPQEAFLPIIIG-----QPDQGNVSTWGF--PF 545
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 BEIDFQLLVLPNIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHS----- 224
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 546 EKASYGTVVYKWLNGANAMKLSFGQSIGIPITNVTSGEYQIRCRYASNDNTNVFFNVD 605
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 -----AQLNPFVHFAKSFYDDIFTGTWIAQKLTLLIAQ----- 259
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 606 GGANPIFQQINFASVDNNTGVQANGVYVYKSIATTDNSFTETPAKTINVHLTNQSSD 665
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 TWQKKQELPSNGSLPYH-VYNIKSIQVTSKVSFSSRQDHSKWCVSIKGSANRWTCIGDL 318
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 666 VELDRIEFIPFSLPIYHGSYNTSS-GADDVLWSSSNMNY--YDIIVNGQANS-----SSI 718
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 NRSLSHQLRG 328
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 719 ASSMHLINKG 728
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 15
US-08-304-626-2
; Sequence 2, Application US/08304626
; Patent No. 5616495
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Kennedy, M. Keith
; APPLICANT: Randall, John Brooks
; APPLICANT: Meier, Henry
; APPLICANT: Uick, Heidi Jane
; APPLICANT: Poncerrada, Luis
; APPLICANT: Schnepf, Harry B.
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: No. 5616495el Bacillus thuringiensis Isolates
; TITLE OF INVENTION: Active Against Hymenopteran Pests and Genes Encoding
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,626
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/887,980
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/SCJ 104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; STRAIN: PS17
; INDIVIDUAL ISOLATE: PS17a
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC1627) NRRL B-18651
; US-08-304-626-2

Query Match      4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97; Indels 112; Gaps 16;
Matches 76; Conservative 48; Mismatches 134;

Qy 31 EYGEADVDFIFKLPKRTSKASEAGLQYLYLDSTR-----QTWNK----- 71
Db 399 EVQDSVETRLYGQLP-----AVDPQAGENYYSIDSSNFIIQINMDTWKTPPQGASGWNNTL 454

Qy 72 -----SYLINSTRSALGR-----TLQHLVDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVSGLSFLORDGTRLSAGMGGFADTIYSLPATH-----YLSYLYGTPYQTSNDY 507

Qy 116 SRQYGHAKGLLVNRTQGFLLHSPK---PPVHGYEYPTSGRRYGTGICITFGYSQF 172
Db 508 SGHVGALVGV-----STPORATLPNTIG-----QPDEQGNVSTMGP-PF 545

Qy 173 HEIDFQLLVLPNTYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHS----- 224
Db 546 EKASYGGTVVKEWLNGANAMKLSFGQSIGIPITNVTSGEYQIRCKYASNDNTNVFENVDT 605

Qy 225 -----AQLNFVHFPAKSFYDDIPTGWIQAKLTHLLAQ----- 259
Db 606 GGANPIFQIINFASVDNNTGVQANGVYVVKSIATDINSFTEIPAKTINVHLTNQSSD 665

Qy 260 TWQKKQELPSNCSLPYH-VVNIKISIGVTSKSYFSSRODHSKWCVKSIKGSANRWTCIGDL 318
Db 666 VFLDRIEFIPFSLPLIVHGSYNTSS-GADDVWSSSNMY--YDIIVNGQANS-----SSI 718

Qy 319 NRSLHQAALRG 328
Db 719 ASSMHLNKG 728
```

Search completed: October 14, 2004, 00:33:35
Job time : 33.3671 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:41:52 ; Search time 2984.85 Seconds
(without alignments)
12245.607 Million cell updates/sec

Title: US-10-790-589-1

Perfect score: 1224

Sequence: 1 tccagtcgccatgga.....aacgtagcgtccataaag 1224

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1216.6	99.4	2689	11 AK079224	AK079224 Mus muscu
2	595.4	48.6	695	14 CF725003	CF725003 UI-M-G20-
3	510	41.7	510	9 A1526873	A1526873 uj47d10.x
4	485.2	39.6	893	13 BQ681950	BQ681950 AGENCOURT

RESULT 1

AK079224

LOCUS

DEFINITION

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE PUBMED	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
REFERENCE AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL REFERENCE AUTHORS	5 Nature 409, 685-690 (2001) The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL REFERENCE AUTHORS	6 Nature 420, 563-573 (2002) (bases 1 to 2689) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE JOURNAL	Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.
FEATURES source	Location/Qualifiers 1..2689 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:9530041E10" /db_xref="MGI:239700" /db_xref="taxon:10090" /clone="9530041E10" /sex="male" /tissue type="urinary bladder" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev stage="adult" 335..1399 /note="unnamed protein product; deoxyribonuclease II beta putative" (MGI:1913283, GB NM_019957, evidence: BLASTN, 99%, match=1529) /codon_start=1 /protein_id="BAC37579.1"
CDS	

db_xref="GI:26347861"	
/translation="MTAKPLRTVLSLLFALLSGVLGTPEISCRNEYGEAVDFWIFYKL PRTSKASREAGLOVLDSTRTWNKSLYLINLSALCKTLOHLYDTHNNTDTAY LIYDQVPGSVNYRQYGHAKLLVNRTOGFWLHSPKPPVHYEYPTSGRGYGO TACITFTGFSQFEEDFQLLVLPQNPISCFISTFHWKLIYMPRCANSSSLKIPVY LAEHSAOGLNFVHFASKFSFYDDIFTGWIAOKLKHLLAQLTQWKKQBLPNCSLPY HYVNIKSIQVTSKSYFSSRQDSKWCVSIGKSANRWTCIGDLNRSJHQAIRGGGFICT KNHIYQAFPHKLYLRGFC"	
ORIGIN	
Query Match	99.4%; Score 1216.6; DB 11; Length 2689;
Best Local Similarity	99.7%; Pred. No. 0;
Matches 1219; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	2 CCAGTCCCTCATGGAATGAAGCCACAGATAGAAAATGACAGCAAAAGCCCTCTAAGAA 61
DB	
QY	297 CACAGTCCCTCATGGAATGAAGCCACAGATAGAAAATGACAGCAAAAGCCCTCTAAGAA 356
DB	
QY	62 CAGTTCCTTTTGTCTCTTTTGGCCCTCTCTGGGGTCTCTGGGGACACCAAGAAATCTCAT 121
DB	
QY	357 CAGTTCCTTTTGTCTCTTTTGGCCCTCTCTGGGGTCTCTGGGGACACCAAGAAATCTCAT 416
DB	
QY	122 GCAGAAATGAATGATGTAAGCTGTGGACTGGTTTATCTTTTATAAGTTACCCAAAGGA 181
DB	
QY	417 GCAGAAATGAATGATGTAAGCTGTGGACTGGTTTATCTTTTATAAGTTACCCAAAGGA 476
DB	
QY	182 CTAGCAAGCACTGAAGAGGGGGGCTGCAGTACCTGTACCTGGACTCCCAAGACAAA 241
DB	
QY	477 CTAGCAAGCACTGAAGAGGGGGGCTGCAGTACCTGTACCTGGACTCCCAAGACAAA 536
DB	
QY	242 CTTGGAACCAAGAGCCCTCTACCTGATTAACAGACCAAGAGTGTCTGGGGAGGACCTTAC 301
DB	
QY	537 CTTGGAACCAAGAGCCCTCTACCTGATTAACAGACCAAGAGTGTCTGGGGAGGACCTTAC 596
DB	
QY	302 AGCATCTGTATGACACACATAATTTCCAGATCAACAGACAGCAGCTATCTAATATACAACGATG 361
DB	
QY	597 AGCATCTGTATGACACACATAATTTCCAGATCAACAGACAGCAGCTATCTAATATACAACGATG 656
DB	
QY	362 GTGTCCTCTGATCTGTGAATTAACAGACAGATGATGACATGCAAGAGTCTGCTGTAT 421
DB	
QY	657 GTGTCCTCTGATCTGTGAATTAACAGACAGATGATGACATGCAAGAGTCTGCTGTAT 716
DB	
QY	422 GGAACAGAACCGAGGGGTCTGGCTGATACACTCTGTCTCCCAAGTTTCCCGCAGTTCATG 481
DB	
QY	717 GGAACAGAACCGAGGGGTCTGGCTGATACACTCTGTCTCCCAAGTTTCCCGCAGTTCATG 776
DB	
QY	482 GCTATGATGATCCCAACCTCCGGGAGGGGATATGACAAACCGGCATCTGCATCACTTTCG 541
DB	
QY	777 GCTATGATGATCCCAACCTCCGGGAGGGGATATGACAAACCGGCATCTGCATCACTTTCG 836
DB	
QY	542 GATACAGCCAGTTTGAAGAAATAGATTTTTCAGCTCTTTGGTCTTACAACCAACATCTACA 601
DB	
QY	837 GATACAGCCAGTTTGAAGAAATAGATTTTTCAGCTCTTTGGTCTTACAACCAACATCTACA 896
DB	
QY	602 GCTCTTCATTTCCAAAGCACTTTTCACTGGAAAATTATCTATCATGCCCCGGAGTGTGCA 661
DB	
QY	897 GCTCTTCATTTCCAAAGCACTTTTCACTGGAAAATTATCTATCATGCCCCGGAGTGTGCA 956
DB	
QY	662 ACTCAGTTCCTTAAAGATCCCTGTCGGTACCTGCTGAACTGCACTCAGCCACAGGGTTC 721
DB	
QY	957 ACTCAGTTCCTTAAAGATCCCTGTCGGTACCTGCTGAACTGCACTCAGCCACAGGGTTC 1016
DB	
QY	722 TAAACTTCGTCCTCAATTTTGGCAAAATCAAGTTTATCTGATGACATCTTTACAGATGGA 781
DB	
QY	1017 TAAACTTCGTCCTCAATTTTGGCAAAATCAAGTTTATCTGATGACATCTTTACAGATGGA 1076
DB	
QY	782 TAGCTCAAAAGTTGAAGACACATTTTGTAGCAGAAAACCTGGCAGAAAAGAAACAAGAGC 841
DB	
QY	1077 TAGCTCAAAAGTTGAAGACACATTTTGTAGCAGAAAACCTGGCAGAAAAGAAACAAGAGC 1136
DB	
QY	842 TTCTCTCAAACTGTGTTCCCTGCTTACCATGCTCTACCAATCAAGTCCATTGGGGTAACCT 901
DB	
QY	1137 TTCTCTCAAACTGTGTTCCCTGCTTACCATGCTCTACCAATCAAGTCCATTGGGGTAACCT 1196
DB	

902 CCAAGTCTTACTTCACTTCTGCGAAGACCATTCCTCAATGCTGTGTTTCCATAAAGGGCT 961
 1197 CCAAGTCTTACTTCACTTCTGCGAAGACCATTCCTCAATGCTGTGTTTCCATAAAGGGCT 1256
 962 CGCGAATCGCTGGACCTGATGAGACCTAAATCGAAGCCCTACACCAAGCCTTAAAGAG 1021
 1257 CGCGAATCGCTGGACCTGATGAGACCTAAATCGAAGCCCTACACCAAGCCTTAAAGAG 1316
 1022 GTGGAGGATTCATCTGTACAAAGATCACTATATACAGGATTCATTAATATATC 1081
 1317 GTGGAGGATTCATCTGTACAAAGATCACTATATACAGGATTCATTAATATATC 1376
 1082 TCCGTTATGGTCTCTGTAAGTAACTCGGTGAAGGCCACACCTCTGTCCCTTGAAGACA 1141
 1377 TCCGTTATGGTCTCTGTAAGTAACTCGGTGAAGGCCACACCTCTGTCCCTTGAAGACA 1436
 1142 GTGGACCTGGAACATCTGCGCTTGGATCTGTTCTCATTAATATCAAGGCTTCTGAGTGAG 1201
 1437 GTGGACCTGGAACATCTGCGCTTGGATCTGTTCTCATTAATATCAAGGCTTCTGAGTGAG 1496
 1202 CACAACGTAGCTCCCAATAAAG 1224
 1497 CACAACGTAGCTCCCAATAAAG 1519

RESULT 2
 LOCUS CF725003
 DEFINITION UI-M-G20-cjo-i-11-0-UI.r1 NIH_BMAP_G20 Mus musculus cDNA clone
 IMAGE:30606058 5', mRNA sequence.
 ACCESSION CF725003
 VERSION CF725003.1 GI:37599171
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. .695
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30606058"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_G20"
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail
 is TTAATGAAGT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 48.6%; Score 595.4; DB 14; Length 695;
 Best Local Similarity 99.2%; Pred. No. 2.3e-175;
 Matches 640; Conservative 0; Mismatches 1; Indels 4; Gaps 2;

QY 2 CCAGTCCCTCCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 61
 DB 52 CACAGTCCCTCCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 111
 QY 62 CAGTCTTTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACAGAAATCTCAT 121
 DB 112 CAGTCTTTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACAGAAATCTCAT 171
 QY 122 GCAGAAATGAATATATGATGAGCTGTGAGCTGTGATCTTTTATTAAGTTTACCCAAAGGA 181
 DB 172 GCAGAAATGAATATGATGAGCTGTGAGCTGTGATCTTTTATTAAGTTTACCCAAAGGA 231
 QY 182 CTAGCAAGGCAAGTGAAGAGGGGGCTGACGTACCTGTACCTGACCTCCCAAGACAA 241
 DB 232 CTAGCAAGGCAAGTGAAGAGGGGGCTGACGTACCTGTACCTGACCTCCCAAGACAA 291
 QY 242 COTGGAACAAGAGCCTCTACCTGATTAACAGCAGGAGTCTCTGGGAGGACCTTAC 301
 DB 292 COTGGAACAAGAGCCTCTACCTGATTAACAGCAGGAGTCTCTGGGAGGACCTTAC 351
 QY 302 AGCATCTGTATGACACACATTAATCCAGATGACACAGCCTTCTAATATACACGATG 361
 DB 352 AGCATCTGTATGACACACATTAATCCAGATGACACAGCCTTCTAATATACACGATG 411
 QY 362 GTGTCCCTGGATCTGTGAATTAACAGCAGATGACATGCCCAAGGTCCTGGTAT 421
 DB 412 GTGTCCCTGGATCTGTGAATTAACAGCAGATGACATGCCCAAGGTCCTGGTAT 471
 QY 422 GGAACAGAACCCAGGGGTTCTGGCTGTATACACTCTGTTCCTCAAGTTTCCCAAGTTTCA 481
 DB 472 GGAACAGAACCCAGGGGTTCTGGCTGTATACACTCTGTTCCTCAAGTTTCCCAAGTTTCA 531
 QY 482 GCTATGATGACCCAACTCCGGG-GAGGATATGACAAACCGGATCTGATCACT-TT 539
 DB 532 GCTATGATGACCCAACTCCGGGAGGCGATATGACAAACCGGATCTGATCACTTTT 591
 QY 540 CGGATACAGCCAGTTT-GAGGAATAGATTTTCAGCTCTTGGTCTTCTTACCAACCAATCT 598
 DB 592 CGGATACAGCCAGTTTGGAGGAATAGATTTTCAGCTCTTGGTCTTCTTACCAACCAATCT 651
 QY 599 ACAGCTGCTTCATTTCCAGACACCTTTTCACTGGAACTTATCTACA 643
 DB 652 ACAGCTGCTTCATTTCCAGACCACTTTTCACTGGAACTTATCTACA 695

RESULT 3

AI526873/c
 LOCUS AI526873
 DEFINITION u147d10.x1 Sugano mouse liver m1ia Mus musculus cDNA clone
 IMAGE:1923091 3' similar to FR:000115 000115 HYPOTHETICAL HUMAN
 PROTEIN R31240_2.; mRNA sequence.
 ACCESSION AI526873
 VERSION AI526873.1 GI:4441008
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 510)
 Marra.M., Hallier.L., Kucaba.T., Martin.J., Beck.C., Wylie.T.,
 Underwood.K., Steptoe.M., Theising.B., Allen.M., Bowers.Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, F., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:979383
Seq primer: custom primer used
High quality sequence stop: 490.

FEATURES

source

1. .510
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1923091"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/notes="Organ: liver; Vector: pME188-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CAACATGTG); 1st strand cDNA was primed with an oligo (dT) primer
[ATGTGGCTCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCACTGG], digested and cloned into distinct DraIII sites of the pME188-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGTCTTAAAGCTGGG and 3' end primer CGACCTGCAGCTCGAGCAC."

ORIGIN

Query Match 41.7%; Score 510; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.5e-148; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0;

715 CAGGGCTAAACCTTCGTCATTTTGCATAATCAAGTTTATATCTGATGACATCTTACA 774
|||||
510 CAGGGCTAAACCTTCGTCATTTTGCATAATCAAGTTTATATCTGATGACATCTTACA 451
|||||

775 GGATGGATAGCTCAAAAGTTGAAGACACATTTGTAGACAAACCTGGCGAAAGAAA 834
|||||
450 GGATGGATAGCTCAAAAGTTGAAGACACATTTGTAGACAAACCTGGCGAAAGAAA 391
|||||

835 CAAGAGCTTCCTTCAAACTGTTCCTGCTTACCATGTCTACACATCAAGTCCATTTGGG 894
|||||

390 CAAGAGCTTCCTTCAAACTGTTCCTGCTTACCATGTCTACACATCAAGTCCATTTGGG 331
|||||

895 GTAACTTCCAAAGTCTTACTTCAGTTCTCGCGAAGCCATTTCCAAATGGTGTTCATA 954
|||||

330 GTAACTTCCAAAGTCTTACTTCAGTTCTCGCGAAGCCATTTCCAAATGGTGTTCATA 271
|||||

955 AAGGGCTCCGCAATCCGCTGGACCTGATGGAGACCTTAATCGAAGCTACCAAGCC 1014
|||||

270 AAGGGCTCCGCAATCCGCTGGACCTGATGGAGACCTTAATCGAAGCTACCAAGCC 211
|||||

1015 TTAAGAGGTGAGGATTCATCTGTACAAAGATCACTACATTTACCAGGCAATTTCAAAA 1074
|||||

210 TTAAGAGGTGAGGATTCATCTGTACAAAGATCACTACATTTACCAGGCAATTTCAAAA 151
|||||

1075 TTATATCTCCGTTATGGGTTCTGTAAGTAACTCGTGGAAGGGCACACCTCTGTCTTT 1134
|||||

150 TTATATCTCCGTTATGGGTTCTGTAAGTAACTCGTGGAAGGGCACACCTCTGTCTTT 91
|||||

1135 GAAACACTGGCACTGGACATCTCGCCTTGGATCTGTCTCCATAATATCAAGGCTTCT 1194
90 GAAACACTGGCACTGGACATCTCGCCTTGGATCTGTCTCCATAATATCAAGGCTTCT 31
1195 GAGTGAGCAACAACGCTAGCGTCCAAATAAAG 1224
30 GAGTGAGCAACAACGCTAGCGTCCAAATAAAG 1

RESULT 4
BQ681950 893 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8183335 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262400
DEFINITION 5', mRNA Sequence.
ACCESSION BQ681950
VERSION BQ681950.1 GI:21794629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 893)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLC2424 row: k column: 09
High quality sequence stop: 647.

FEATURES
source
1. .893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6262400"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

ORIGIN
Query Match 39.6%; Score 485.2; DB 13; Length 893;
Best Local Similarity 74.7%; Pred. No. 1.2e-140;
Matches 623; Conservative 0; Mismatches 205; Indels 6; Gaps 1;

356 ACAGTGTGTCCTCGGATCTGTGAATTCAGCAGACAGTATGGACATGCCAAAGTCTGC 415
|||||
1 ATGATGGAGTCCCTTAAACCTGTGAATTCAGCAGACAGTATGGACATGCCAAAGTCTGC 60
|||||

416 TGGTATGCACAGAACCCAGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCCCAG 475
|||||

61 TCGTGTGGAACAGAGTTCAAGGGTTCTGGCTGATTCATTCATCCCTCAGTTTCTCCAA 120
|||||

476 TTC-----ATGGCTATGAGTACCCAACTCGGGAGCGGATATGACAAACCGGATCT 529
|||||

121 TTCGGGAAGAAGCTATGATTATCCACCACAGGAGACGAAATGGACAAAGTGGCATCT 180
|||||

530 GCATCACTTCGGATACAGCCAGTTTGAGGNAATAGATTTTCAGCTCTTGCTTTACAAC 589
|||||

181	DB	GCATTAACCTTTCAAGTACAAACAGATTAGAGCAATAGATTCCTAGCTCTTGCTGTGCACCC	240
590	QY	CAAAACATCTACAGCTGCTTCATTCCAAAGCACCTTTCACTGGAAACTTATCTACATGCCCC	649
241	DB	CCAAAGCTCTATAGCTGCTCCATCCCAAGCCACTTTTCCACGAGAGTCATTCACATGCCCC	300
650	QY	GGATGTGCGCCAACTCCAGTTCCTTAAGATGCCCTGTCCGGTACCTCGCTGAATGCACTGACT	709
301	DB	AGCTGTGCACACGGCCAGCTCATCAGAGATTCCTGGCAGGCTCTCTCACCACTCTTCAGT	360
710	QY	CAGCCCAAGGCTTAAACTTCGTCATTTTGCAAAATCAAAGTTTTTATCTACATGACATCT	769
361	DB	CGGCCCAAGGCAAAAATTCCTCATTTTGCAAAGTCGGATTCCTTTCTTGAACACATCT	420
770	QY	TTACAGGATGGATAGCTCAAAAGTTGAAGACACATTTGTGTAGCACAAACCTGGCAGAAA	829
421	DB	TTGACAGCCTGGATGGCTCAACGGCTGAAGACACACTTGTTAACAGAAACCTGGCAGCGAA	480
830	QY	AGAAACAAAGAGCTTCCTTCAAACCTGTTCCCTGCCTTACCATGCTCAACAACATCAAGTCCA	889
481	DB	AAAGACAAGAGCTTCCTTCAAACCTGCTCCCTTCCTTACCATGCTCAATATAAAGCAA	540
890	QY	TTGGGGTAACTTCCAAAGCTTACTTTCAGTTCCTGCCAAGACCATTCCAAATGGTGTGTTT	949
541	DB	TTAAATATCACGACACTCTATTTCAGTTCCTTATCAAGATCAACGCAAGTGTGTATTT	600
950	QY	CCATAAAGGGCTCCGCAAACTCGCTGGACCTTGCAATGGAGACCTAAATCGAAGCCCTACACC	1009
601	DB	CCCAAAAGGGCACCAAAAATCGCTGGACATGATTTGGAGACCTTANATCGGAGTCCACACC	660
1010	QY	AAGCCTTAAGAGTGGAGGATTCACTGTGTCAAGAAGATCACTACATTTACAGGCATTTTC	1069
661	DB	AAGCCTTCAAGAGTGGAGGATTCATTTGTGTACCCAGAAATGGCAAAATTTACCAAGCATTTN	720
1070	QY	ATAAAATTATATCTCCGTTATGGGTTCTGTAAAGTAAACTCGCTGAAGAGCCACACCCCTCTG	1129
721	DB	CAGGATTAGTATTATACTATGAAGCTGTAGTAAACTTCGTGAAGAGCACAGGTACTA	780
1130	QY	TCCTTGAAGAACACTGGCACTGGGAACATCTCGCCTGGATCTGTTCTCCATAATA	1183
781	DB	TCAATTGAAGACCTTGAACATGGGTCTCTTCCATTAACCTTCTTTATATTTTA	834

```

RESULT 5
BX526247
BX526247
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX526247 546 bp mRNA linear EST 27-JUN-2003
BX526247 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE998M204718 ; IMAGE:1923643, mRNA sequence.
BX526247
BX526247.1 GI:32297078
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Heil,O., Ebert,J., Neubert,P., Peters,M., Radelof,U., Schneider,D.
and Korn,B.
Mouse UnigeneSet - RZPD2
Unpublished (2003)
Contact: Ina Rolf
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE998M204718.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi?responseLibNo=981 Contact: Ina Rolf
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111

```

www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
sugF, Primer sequence: CTTTGTCTAAAGAGCTGG.

FEATURES

```

1..546
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGp98M204718 ; IMAGE:1923643"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site.1: Drail1
(CACGTGTGG); Site 2: Drail1 (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCGCTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a Drail1 adaptor [TGTGGCCTACTGG], digested
and cloned into distinct Drail1 sites of the pME18S-FL3
vector (5' site CACGTGTGG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTCTCTAAAAGCTGG and 3' end
primer CGACCTGTGACCTGCAGCACA."

```

ORIGIN

Query Match	37.5%;	Score 459.4;	DB 13;	Length 546;
Best Local Similarity	99.8%;	Pred. No. 1.3e-132;		
Matches 460;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy 2	CCCAAGTCCCTGCATGTGAATGAAGGCCACAGATAGAAAAATGACAGCAAAAGCCTCTTAAGAA	61		
Db				
86	CACAGTCCCTGCATGTGAATGAAGGCCACAGATAGAAAAATGACAGCAAAAGCCTCTTAAGAA	145		
Qy 62	CAGTTCTTTCTTTTGCTCTTCTTTTGCCCTCTCTGGGTCCTCTGGGGACACACAGAATCTCAT	121		
Db				
146	CAGTTCTTTCTTTTGCTCTTCTTTTGCCCTCTCTGGGTCCTCTGGGGACACACAGAATCTCAT	205		
Qy 122	GCAGAAATGAATATGTTGAAGCTGTGGACTGTGGTTTATCTTTTATAAGTTACCCCAAAGGA	181		
Db				
206	GCAGAAATGAATATGTTGAAGCTGTGGACTGGTTTATCTTTTATAAGTTACCCCAAAGGA	265		
Qy 182	CTAGCAAGGCAAGTGAAGAGCGGGGCTGCAGTACCTGTACCTTGGACTCCACAAGACAAA	241		
Db				
266	CTAGCAAGGCAAGTGAAGAGCGGGGCTGCAGTACCTGTACCTTGGACTCCACAAGACAAA	325		
Qy 242	CTTGGACACAGAGCCTCTACCTGATTTAAACAGCACAGGAGTGCTCTGGGGAGGACCTTAC	301		
Db				
326	CTTGGACACAGAGCCTCTACCTGATTTAAACAGCACAGGAGTGCTCTGGGGAGGACCTTAC	385		
Qy 302	AGCATCTGTATGACACACATAATTTCCACGAATGACACAGCCTATCTAATATACACAGATG	361		
Db				
386	AGCATCTGTATGACACACATAATTTCCAGGAATGACACAGCCTATCTAATATACACAGATG	445		
Qy 362	GTGTCCCTGGATCTGTGAATTTACAGCAGACAGTATGGAATGCCAAAGCTCTGTGGTAT	421		
Db				
446	GTGTCCCTGGATCTGTGAATTTACAGCAGACAGTATGGAATGCCAAAGCTCTGTGGTAT	505		
Qy 422	GGNACAGAGCGAGGGGTTCTGGCTGATACATCTGTGTTCCC	462		
Db 506	GGNACAGAGCGAGGGGTTCTGGCTGATACATCTGTGTTCCC	546		

RESULT 6	AI875954	432 bp	mRNA	linear	EST 21-JUL-1999
LOCUS	AI875954				
DEFINITION	uj5b10.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1923643 5', similar to TR:090759 DEOXYRIBONUCLEASE II', mRNA sequence.				

QY	242	CCTCGAACAAGAGGCTCTACCTGATTAACAGACACAGAGTGTCTCTGGGAGGACCTTAC	301
Db	260	CCTGGAAACAAGAGCTCTACCTGATTAACAGACACAGAGTGTCTCTGGGAGGACCTTAC	319
QY	302	AGCATCTGTATGACACACATATTCACGAATGACACAGCCTATCTAATATACACAGATG	361
Db	320	AGCATCTGTATGACACACATATTCACGAATGACACAGCCTATCTAATATACACAGATG	379
QY	362	GTGTCCCTCGATCTGTGAATTACAGACACAGATGACATGCCAAAGGCTCTG	414
Db	380	GTGTCCCTCGATCTGTGAATTACAGACACAGATGACATGCCAAAGGCTCTG	432
RESULT 7			
LOCUS	CD365871/c	758 bp	mRNA linear EST 29-MAY-2003
DEFINITION	UI-H-FT2-bjk-p-12-0-UI.s1 NCI CGAP_F12 Homo sapiens cDNA clone		
ACCESSION	UI-H-FT2-bjk-p-12-0-UI 3', mRNA sequence.		
VERSION	CD365871		
KEYWORDS	CD365871.1 GI:31149961		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 758)		
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
FEATURES	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Dr. Gary W. Humminghake, U of I		
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Distribution information can be found at		
ORIGIN	http://genome.uiowa.edu/distribution/cgap.html		
	Seq primer: M13 FORWARD		
	POLYA=Yes.		
	Location/Qualifiers		
	1..758		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="UI-H-FT2-bjk-p-12-0-UI"		
	/tissue type="Aveolar Macrophage"		
FEATURES	/dev stage="Adult"		
	/lab_host="DH10B (Life Technologies)"		
	/clone_lib="NCI_CGAP_F12"		
	/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a		
	modified polylinker; Site 1: EcoR I; Site 2: Not I;		
	NCI CGAP_F12 is a subcloned cDNA library constructed from		
	a pool of 81 RNA samples from Alveolar Macrophages		
	challenged with different treatments. The library was		
	subtracted according to Bonaldo, Lennon and Soares, Genom		
	Research, 6:791-806, 1996. The tissue was provided by Dr.		
ORIGIN	Gary W. Humminghake of the University of Iowa.		
	TAG TISSUE=Human Lung Aveolar Macrophage		
	TAG LIB=UI-H-FT2		
	TAG_SEQ=GGCCATGCGC		
	Query Match 32.7%; Score 400; DB 14; Length 758;		
	Best Local Similarity 75.0%; Pred. No. 6.8e-114;		
	Matches 513; Conservative 0; Mismatches 170; Indels 1; Gaps 1		
QY	500	CGGGAGCGCATATGGACAAACCGCATCTGCATCACTTTCCGATACAGCAGTTTGAAG	559
Db	757	CAGGAGCAGCAATGGACAAA-GTGCATCTGCATCACTTTCAAGTACACACAGTATGAG	699
QY	560	AAATAGATTTTCAGCTTTGGTCTTCAACCAACCACTACAGCTGCTTCATTCAGCA	619

Db 698 CAATAGATTCTCTAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGCTCCATCCAGCCA 639
 QY 620 CCTTTCACCTGAAACCTTATCTACATGCCCGGATGTGCACTCCAGTCTCTTAAGA 679
 Db 638 CCTTTCACAGGAGCTCATTCATGCCCCAGCTGTGCACAGGCGCAGCTCATCAGAGA 579
 QY 680 TCCTGTCGCGTACTCTGCTGTAATCTGCACTCAGCCAGGGTCTAAATCTTGCTCCATTTG 739
 Db 578 TTCTGTCGAGCTCTTCAACACACTTCAGTCGCGCCAGGCAAAAATCTCTCCATTTG 519
 QY 740 CAAATCAAGTCTTTTATCTAGTACATCTTTTACAGATGATAGTCAAAAATCTGAAGA 799
 Db 518 CAAAGTCGAGTCTTTTCTTGTGATGATCTTTTGCAGCTGATGCTCAACGCTGAAGA 459
 QY 800 CACATTTGTTAGCACAAACCTGCGACAAAAGAAACAAAGAGCTTCTTCAAACTGTTCCC 859
 Db 458 CACACITGTTTAAACAGAAACCTGCGACGGAAGAAACAGAGCTTCTTCAAACTGTTCCC 399
 QY 860 TGCCTTACCATGCTCTCAACATCAAGTCCATTTGGGTAATCTTCAAGTCTTACTTCAGTT 919
 Db 398 TTCTCTTACCATGCTCTCAATATAAAGCAATTAATATCAAGACACTCTTATTTTCAGTT 339
 QY 920 CTCGCCAAGACCATTCCTCAATAGTGTGTTTCCATTAAGGGCTCGCAAAATCGCTGACCT 979
 Db 338 CTATCAAGATCATGCGCAAGTGTGTTTTCCTCAAAAGGGCACCAAAAATCGCTGACAT 279
 QY 980 GCATTTGGAGACCTTAATCGAAGCCTTACACCAAGCCTTAAGAGGTGAGGATTCATCTGA 1039
 Db 278 GTATTTGGAGACCTTAATCGAGTCTCACACCAAGCCTTCAAGTGGAGGATTCATTTGA 219
 QY 1040 CAAAGATCACTACATTTACCGAGCTTTCATAAATATATCTCTCGTTATGGGTTCTGTA 1099
 Db 218 CCCAGATTTGGCAAAATTTACCAAGCAATTTCAATGATTATTAATATGAAGCTGA 159
 QY 1100 ACTAACTCGGTGAAGGCCACACCTCTGCTTGAAGAACTGCGACATCGGAATCTC 1159
 Db 158 AGTAACTTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACATGGGCTCTTCT 99
 QY 1160 GCCTTGGATCTGTTCTCCATATA 1183
 Db 98 CCATTTACACCTCTTTATATTITA 75

AI048641 543 bp mRNA linear EST 08-JUL-1998
 IMAGE:1450676.5' similar to TR:000115 000115 HYPOTHETICAL HUMAN
 PROTEIN R31240_2. ; mRNA sequence.
 AI048641
 AI048641.1 GI:3296928
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 543)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

TITLE
 JOURNAL
 COMMENT

MGI:923992
 Seq primer: custom primer used
 High quality sequence stop: 518.
 Location/Qualifiers
 1. 543
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1450676"
 /sex="female"
 /dev stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
 (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
 was primed with an oligo(dT) primer
 [ATGTCCTCTTTTCTTTTCTTTT]; double-stranded cDNA was
 ligated to a DraIII adaptor [GTGTGCTTACTGG], digested
 and cloned into distinct DraIII sites of the pME18S-FL3
 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
 be used to isolate the cDNA insert. Size selection was
 performed to exclude fragments <1.5kb. Library
 constructed by Dr. Sumio Sugano (University of Tokyo
 Institute of Medical Science). Custom primers for
 sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end
 primer CGACCTGAGCTCGAGACA."

FEATURES
 source

Query Match 31.2%; Score 382.4; DB 9; Length 543;
 Best Local Similarity 99.7%; Pred. No. 2.1e-108;
 Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCAGTCCCTCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAAAGCTCTTAAGAA 61
 Db 160 CACAGTCCCTCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAAAGCTCTTAAGAA 219
 QY 62 CAGTCTTTCTTTGCTCTTTCTTTGCTCTCTCTGCTGCTCTGCTGCTCTGCTGCTCT 121
 Db 220 CAGTCTTTCTTTGCTCTTTCTTTGCTCTCTCTGCTGCTCTGCTGCTCTGCTGCTCT 279
 QY 122 GCAGAAATGAATATGATGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
 Db 280 GCAGAAATGAATATGATGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339
 QY 182 CTAGCAAGCGCAAGTGAAGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 Db 340 CTAGCAAGCGCAAGTGAAGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
 QY 242 CCTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGGCTGCTGCTGCTGCTGCTGCT 301
 Db 400 CCTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGGCTGCTGCTGCTGCTGCTGCT 459
 QY 302 AGCATCTCTATGACACACATTAATTCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCT 361
 Db 460 AGCATCTCTATGACACACATTAATTCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCT 519
 QY 362 GTGTCCTCTGATCTGTGAATTTACA 385
 Db 520 GTGTCCTCTGATCTGTGAATTTACA 543

ORIGIN

Query Match 31.2%; Score 382.4; DB 9; Length 543;
 Best Local Similarity 99.7%; Pred. No. 2.1e-108;
 Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCAGTCCCTCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAAAGCTCTTAAGAA 61
 Db 160 CACAGTCCCTCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAAAGCTCTTAAGAA 219
 QY 62 CAGTCTTTCTTTGCTCTTTCTTTGCTCTCTCTGCTGCTCTGCTGCTCTGCTGCTCT 121
 Db 220 CAGTCTTTCTTTGCTCTTTCTTTGCTCTCTCTGCTGCTCTGCTGCTCTGCTGCTCT 279
 QY 122 GCAGAAATGAATATGATGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
 Db 280 GCAGAAATGAATATGATGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339
 QY 182 CTAGCAAGCGCAAGTGAAGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 Db 340 CTAGCAAGCGCAAGTGAAGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
 QY 242 CCTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGGCTGCTGCTGCTGCTGCTGCT 301
 Db 400 CCTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGGCTGCTGCTGCTGCTGCTGCT 459
 QY 302 AGCATCTCTATGACACACATTAATTCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCT 361
 Db 460 AGCATCTCTATGACACACATTAATTCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCT 519
 QY 362 GTGTCCTCTGATCTGTGAATTTACA 385
 Db 520 GTGTCCTCTGATCTGTGAATTTACA 543

RESULT 9
 CD365107/c
 LOCUS
 DEFINITION
 UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP_Ft2 Homo sapiens cDNA clone
 UI-H-PT2-bjn-1-05-0-UI 3', mRNA sequence.

ACCESSION
 CD365107
 VERSION
 CD365107.1
 SOURCE
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CD365107 729 bp mRNA linear EST 29-MAY-2003
 UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP_Ft2 Homo sapiens cDNA clone
 UI-H-PT2-bjn-1-05-0-UI 3', mRNA sequence.

ACCESSION
 CD365107
 VERSION
 CD365107.1
 SOURCE
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CD365107 729 bp mRNA linear EST 29-MAY-2003
 UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP_Ft2 Homo sapiens cDNA clone
 UI-H-PT2-bjn-1-05-0-UI 3', mRNA sequence.

ACCESSION
 CD365107
 VERSION
 CD365107.1
 SOURCE
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 729)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue Procurement: Dr. Gary W. Hunninghake, U of I cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/cgap.html Seq primer: M13 FORWARD polVA=Yes.

```

FEATURES             Location/Qualifiers
     1..729
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-H-FT2-bjn-1-05-0-UI"
     /tissue_type="Aveolar Macrophage"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies)"
     /clone_lib="NCI CGAP FT2"
     /notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: BcoR 1; Site 2: Not 1;
NCI CGAP FT2 is a subcloned cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
subcloned according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. The tissue was provided by Dr.
Gary W. Runninghake of the University of Iowa.
TAG TISSUE=Human Lung
TAG LIB=UI-H-FT2
TAG_SEQ=GGCCATGCG"

```

ORIGIN	Query Match	31.0%;	Score 379.2;	DB 14;	Length 729;
	Best Local Similarity	74.5%;	Pred. No. 2.3e-107;		
	Matches 477;	Conservative 0;	Mismatches 163;	Indels 0;	Gaps 0;
QY	544	TACAGCCAGTTTGAGGAATAATGATTTTCAGCTCTTGGTCTTTACAAACAAACATCTACAGC	603		
DB	727	TACAACAGTATGAGCCAAATAGATTCTCAGCTCTTGGTCTGGCAACCCCAAGCTATATAGC	668		
QY	604	TGCTTCATTCCAAAGCACCTTTTCATCTGGAAAACCTTATCATATGCCCCCGGATGTGTGCCAAC	663		
DB	667	TGCTCCATCCACGCCACCTTTCACAGGAGCTCAATTCAGATGCCAGCTGTGCACCAAGG	608		
QY	664	TCCAGTTCCTTAAAGATCCCTGTGCGGTAACCTCGCTGAATGCACTCAGCCCCAGGGTCTA	723		
DB	607	GCCAGCTCATCAGAGATTCCTGGCAGGCTCTCTCCACACTTCAGTGGGCCAGGGACAA	548		
QY	724	AACTTCGTCACATTTTGCAAAATCAAGTTTTTATCTGATGACATCTTTACAGATGATA	783		
DB	547	AAATTCCTTCCATTTTGCAAGTCGGATCTTTTCTTGACACATCTTTGAGCCTGGATG	488		
QY	784	GCTCAAAAAGTTGAAGACACATTTGTTAGCACAAACCTGGCAGAAAAAGAAAACAAGAGCTT	843		
DB	487	GCTCAACGGCTGAAGACACACTTGTTTACAGAAACCTGGCAGCGAAAAAGACAAGAGCTT	428		
QY	844	CCTTCAAACTGTTCCCTGCCTTACATGCTTACAACATCAAGTCCATTTGGGTAACTTCC	903		
DB	427	CCTTCAAACTGCTCCCTTCTTACCATGCTCAATATAAAGCAATTAATATACGA	368		
QY	904	AAGCTCTACTTCAGTTCGCGCAAGACCAATTCCTAAATGGTGTGTTTCCATAAAGGGTCC	963		
DB	367	CACCTCTAATTCAGTTCCTATCAAGATCATGCCAAGTGGTGTATTTCCAAAGGGCACC	308		
QY	964	GCAATTCGCTGACCTGCATTTGGAGACCTTAATCGAAGCCTACACAAAGCTTTAAGGTT	1023		

Db	307	AAAAATCGGTGACATGATTTGGAGAGCCTAAATCGGAGTCCACACCAAGCCTTTCAGAAGT	248
Qy	1024	GGAGGATTCATCTGTTACAAGAATCACTACATTTACCGGCATTTTCATAAATATATATCTC	1083
Db	247	GGAGGATTCATTTTGTACCCAGAATTCGCAAAATTTTCCAGCATTTCAAGGATTAGTATTA	188
Qy	1084	CGTTATGGGTTTCGTAAAGTAATCACTCGGTGAAGGCCACACCCCTCTGTCCTTGAGAAACACT	1143
Db	187	TACTATGAAGCTGTAAGTAAACTTGGTGAAGGACACAGGTACTCATTCATTTGAAACCTT	128
Qy	1144	GGCACTGGAAACATCTCGCCCTGGATCTGTCTCCATAATA	1183
Db	127	GACAAATGGGCTCTCTTCCATATACACCTCTCTTATATTTTA	88

RESULT 10
 BG761484
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens (human)
 mRNA sequence.
 BG761484.1 GI:14072137
 BG761484
 BG761484.1
 EST.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 923)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LiCML713 row: k column: 23
High quality sequence spot: 845.

```

FEATURES
source
1..923
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4858894"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DHI0B (phage-resistant)"
/clone_lib="NIH MGC 49"
/note="organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN	Query Match	29.9%;	Score 366.4;	DB 12;	Length 923;
	Best Local Similarity	73.3%;	Pred. No. 2.7e-103;		
	Matches 597;	Conservative 0;	Mismatches 191;	Indels 26;	Gaps 9
Qy	358	GATGTGTCCTCGATCTGTGAATTAACAGACAGACAGTATGACATGCCCAAGGTCGTG	417		
Db	1	GATGGAGTCCTTAAACCTGTGAATTAACAGCAAAAGTATGGACACACCAAGGTTTAC	60		
Qy	418	GTATGGACACAGAAACGACGGGGTTCCTGGCTGTGATACACTCTGTGTTCCCAAGTTTCC	477		
Db	61	CTGTGGACACAGAGTTTCAAGGGTTCCTGGCTGATTCATTCCTCCTCAGTTTCTCTCCA	120		

478 C-----ATGCTATAGTACCAACCTCGGAGGCGATATGACAAACCGCATCTGC 531
 121 CCGGAAGAAGGCTATGATTATCCACCACAGGAGGAGCAATGCAATGCACTTGC 180
 532 ATCACTTTGGATACAGCAGCTGTTGAGGAATAG-----ATTTTCAGCTCTTGG 580
 181 ATAACTTTCAAGTACACCACTAGTACGCAATAGGTAATAAATTTCTCAGCTCTTGG 240
 581 TCT-TACAAACCAACATCTACAGCTGCTTCATCCCAAG-CACCTTTCACTGGAACTTAT 638
 241 TCTGCAACCTTAAGCTATAGCTGCTCATCCCTAGCACCCTTTACAGGAGCTCAT 300
 639 CTATATGCCCC--GGATGTGTGCAACTCCAGTTCCTTAAAGATCCCTGTCGGTACCTC 696
 301 TCACATGCCCCATGCTGTGCAGCCAGGCGAGCTCATCAGAGATTCCTGGCAGCTCCIC 360
 697 GCTGAAGTCACTACAGCCAGGCTTAAGCTTCGTCATTTGCAAAATCAAGTTTAT 756
 361 ACCACACTTTCAGTCGCGCCAGGAGCAAAAATTCCTCAATTTGCAAGTCGGATCTTTT 420
 757 ACTGATGACATCTTACAGATGATAGCTTCAAACTGTTCCCTGCTTACCATGCT 874
 421 CTGACGACATCTTTCAGCTGATGCTCAACGCTTCAAGCTTGAAGACACACTTGTAAAGAA 480
 817 ACTGGC--AGAAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTGCTTACCATGCT 874
 481 ACTGGCTAGCAATTAAGACAGAGCTTCTTCAAACTGCTCCCTTCTTACCATGCT 540
 875 ACAACATCAAGTCCATTGGGTAACTTCAAGCTTCACTTCACTTCTGCTCCGCAAGCAAT 934
 541 ACAATAAAGCAATTAATAATATACGACACTCTTATTTCAAGTCTTATCAAGATCATG 600
 935 CCAATATGCTGTGTTCCATAAAGGCTCC-GCAATTCGCTGACCTGCATTTGAGACCTA 993
 601 CCAAGTGTGATTTCCCAAAAGGCAACCAAGATTCGCTGACATGATTTGAGACCTA 660
 994 AATCGAAGCTTACCAAGCTTAAAGAGTGG-AGGATTCATCTGTA-CAAAAGATCACT 1051
 661 AATCGAGTCCACAGAGCTTCAAGTGGAGGATTCATTTGTACCAAGATTCG 720
 1052 ACAATTCAGAGCTTCAATAATATATCTCCGTTATGGGTCTGTAAAGTAACTCGGT 1111
 721 AAATTTACCAAGCAATTTCAAGATTTAGTATTTACTATGAAAGCTGTCAAGTAACTTGT 780
 1112 GAAAGCCACACCTCTGCTCTTGAACACACTGG 1145
 781 GAAAGGCCAGGTACTATCCATTTGAACACCTTG 814

RESULT 11
 CD366106/c
 LOCUS
 DEFINITION
 UI-H-FTL-bjs-1-13-0-UI.s1 NCI CGAP_FTL Homo sapiens cDNA clone
 UI-H-FTL-bjs-1-13-0-UI 3', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgap.html>
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..727
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FTL-bjs-1-13-0-UI"
 /issue_type="Aveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_FTL"
 /note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FTL is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGGCATGGCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
 TAG TISSUE=Human Lung Alveolar Macrophage
 TAG_LIB=UI-H-FTL
 TAG_SEQ=GGCATGGCG"

ORIGIN

Query Match 29.6%; Score 362.2; DB 14; Length 727;
 Best Local Similarity 73.6%; Pred. No. 5.2e-102;
 Matches 460; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
 QY 559 GAAATAGATTTTACGCTCTTGGTCTTCAACCAAAACATCTACAGCTGCTTCAATCCAAAGC 618
 Db 712 GCAATAGATTTTACGCTCTTGGTCTTGGCAACCAAGCTATAGCTGCTTCAATCCAAAGC 653
 QY 619 ACCTTTCTACTGGAATCTTATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
 Db 652 ACCTTTCTACTGGAATCTTATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
 QY 679 ATCCCTGTCCGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
 Db 592 ATTCTGTGAGGCTCTCTCAACACATTCAGTGGCCGAGGACAAAATTCCTCATTTT 533
 QY 739 GCAAAATCAAGTTTATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
 Db 532 GCAAAATCGGATTTCTTTTCTTGAACATCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 473
 QY 799 ACACATTTGTTAGTACAAACCTTGGCAGAAAAGAAAAGAGAGCTTCTTCAAACTGTTCC 858
 Db 472 ACACATTTGTTTAAAGAAACCTTGGCAGGAAAAGAGAGAGCTTCTTCAAACTGTTCC 413
 QY 859 CTGCTTACCATGCTTCAACATCAAGTCCATTTGGGTAACTTCCAGTCTTCACTTCACTTCACT 918
 Db 412 CTTCCTTACCATGCTTCAACATCAAGTCCATTTGGGTAACTTCCAGTCTTCACTTCACTTCACT 353
 QY 919 TCTGCCAAGACCATTTCCAAATGCTGTTTCCATAAAGGCTCCGCAAACTCGTGGACC 978
 Db 352 TCTTATCAAGATCAAGCAAGTGTGTGTTTCCATAAAGGCTCCGCAAACTCGTGGACC 293
 QY 979 TGCATTGGAGACCTTAATCGAAGCTTCAACCAAGCTTTAAGAGTGGAGGATTCATCTGT 1038
 Db 292 TGTATTGGAGACCTTAATCGAAGCTTCAACCAAGCTTTAAGAGTGGAGGATTCATCTGT 233
 QY 1039 ACAAGATCACTACATTTTACAGGCTTTCATTAATATATCTCCGTTATGGGTTCTGT 1098

Soares and M. Fatima Bonaldo. "

QY	919	TCCTGGCAAGACCAATTC	25.0%;	Score 305.8;	DB 12;	Length 609;
Db	319	TCTTATCAAGATCATGCCAAGTGTGTATTTCCAAAGGGCCACCAAAATCGCTGGACA	Best Local Similarity 72.3%;	Pred. No. 2.7e-84;	Indels 0;	Gaps 0;
QY	979	TGCAATGGAGACCTAAATCGAAGCTTACACCAAGCTTAAAGAGTGGAGGATTCATCTGT	Matches 397;	Conservative 0;	Mismatches 152;	
Db	259	TGTAITGGAGACCTAAATCGAGTCCACACCAAGCCTTCAGAAGTGGAGGATTCATTTGT				
QY	1039	ACAAAGATCAGTACATTTACACAGGCAATTTCAFAATTTATATCTCGTTAAGGTTCTGT				
Db	199	ACCCAGATTTGGCAATTTTACCAAGCATTTCAAGGATTTAGTATTATATGAAAGCTGT				
QY	1099	AGTAACTCGTGAAGGCCACACCTCTGTCTTGAAGAACTGGCACTGGAACATCT				
Db	139	AGTAACTTGTGTGAAGGACACAGGTACTATCATTTGAAACCTTGACAAATGGGTCTTCT				
QY	1159	CGCTTTGGATCTCTTCCATATATATCAAGCTTCTGAGTGGACAAACGTAAGCTCCAA				
Db	79	TCATTACACC--TTCTTATATTTTAAAGGCTGTGAATACTTATAACCTGCATATCA				
QY	1219	TAAAA 1223				
Db	21	CAAAA 17				

RESULT 15

EG149286/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Search completed: October 14, 2004, 05:24:09
Job time : 2992.85 secs

Result No.	Query τ			DB	ID	Description
	Score	Match	Length			
1	1224	100.0	1224	93	US-09-949-434-1	Sequence 1
2	1224	100.0	1224	13	US-10-240-709-1	Sequence 1
3	1224	100.0	1224	17	US-10-790-589-1	Sequence 1
4	1219.8	99.7	1652	17	US-10-670-863-2	Sequence 2
5	672.8	55.0	1268	9	US-09-949-434-3	Sequence 3
6	672.8	55.0	1268	13	US-10-240-709-3	Sequence 3
7	672.8	55.0	1268	17	US-10-790-589-3	Sequence 3
8	641.8	52.4	1086	17	US-10-670-863-4	Sequence 4
9	193.2	15.8	480	10	US-09-918-995-7149	Sequence 4
10	65	5.3	65	10	US-09-908-975-28714	Sequence 4
11	60.8	5.0	1575	16	US-10-408-167A-1	Sequence 1
12	53.8	4.4	65	10	US-09-908-975-2418	Sequence 1
13	48.8	4.0	60	10	US-09-908-975-13222	Sequence 1
14	41	3.3	379	9	US-09-960-352-7234	Sequence 7

Query Match	100.0%;	Score 1224;	DB 9;	Length 1224;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1224;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCCAGTCCCTGCATGGAAATGAAGGCCACAGATAGAAATGACAGAAAGCCTTAAGA	60	
Ddb	1	TCCAGTCCCTGCATGGAAATGAAGGCCACAGATAGAAATGACAGAAAGCCTTAAGA	60	
QY	61	ACAGTCTCTTTTGTCTCTCTTTTGGCCCTCTCTGGGGTCTCTGGGGACACAGAAATCTCA	120	
Ddb	61	ACAGTCTCTTTTGTCTCTCTTTTGGCCCTCTCTGGGGTCTCTGGGGACACAGAAATCTCA	120	
QY	121	TGCAGAAATGAATATCGTGAAGCTGTGGACTGGTTATCTTTTATAAGTTACCCAAAAGG	180	
Ddb	121	TGCAGAAATGAATATCGTGAAGCTGTGGACTGGTTATCTTTTATAAGTTACCCAAAAGG	180	

181 ACTAGCAGGCAAGTGAAGAGCGGGCTGCTGAGTACCTGCTGACCTCCACAGCAA 240
181 ACTAGCAGGCAAGTGAAGAGCGGGCTGCTGAGTACCTGCTGACCTCCACAGCAA 240
241 ACCTGGAACAAGCCTCTACCTGATTAAACAGCAGGAGTCTCTGGGAGGACCTTA 300
241 ACCTGGAACAAGCCTCTACCTGATTAAACAGCAGGAGTCTCTGGGAGGACCTTA 300
301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATAATACAGAT 360
301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATAATACAGAT 360
361 GGTGTCCTGATCTGTGATTAACAGACAGATATGGACATGCCAAAGTCTGCTGTA 420
361 GGTGTCCTGATCTGTGATTAACAGACAGATATGGACATGCCAAAGTCTGCTGTA 420
421 TGGACACAGACGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480
421 TGGACACAGACGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480
481 GGCTATGAGTACCAACCTCGGGAGCGGATATGGACAAACCGGCTATCTGCACTTTC 540
481 GGCTATGAGTACCAACCTCGGGAGCGGATATGGACAAACCGGCTATCTGCACTTTC 540
541 GGTATGAGCAGTTTGAAGAAATAGATTTTTCAGTCTTGTGCTTACCAACCAATCTAC 600
541 GGTATGAGCAGTTTGAAGAAATAGATTTTTCAGTCTTGTGCTTACCAACCAATCTAC 600
601 AGCTGCTTCACTTCAAGCAGCTTCTGCTGAACTTATCTACATGCTCCCGGATGTGTC 660
601 AGCTGCTTCACTTCAAGCAGCTTCTGCTGAACTTATCTACATGCTCCCGGATGTGTC 660
661 AACTTCCAGTTTCCCTTAAAGATCCCTGCTCGGTAACCTGCTGAACTGCTGAGGATGG 720
661 AACTTCCAGTTTCCCTTAAAGATCCCTGCTCGGTAACCTGCTGAACTGCTGAGGATGG 720
721 CTAAACTTCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 780
721 CTAAACTTCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 780
781 ATAGTCAAAAGTTGAAGACACATTTGTTAGACAAACCTGGCAGAAAGAAACAAAGAG 840
781 ATAGTCAAAAGTTGAAGACACATTTGTTAGACAAACCTGGCAGAAAGAAACAAAGAG 840
841 CTTCCTTCAAACTGTTCCCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 900
841 CTTCCTTCAAACTGTTCCCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 900
901 TCCAGTCTTACTTCAAGTCTGCTGCAAGACCATTTCCAAATGGTGTGTTCCATAAAGGC 960
901 TCCAGTCTTACTTCAAGTCTGCTGCAAGACCATTTCCAAATGGTGTGTTCCATAAAGGC 960
961 TCCGCAAAATCGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 1020
961 TCCGCAAAATCGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 1020
1021 GGTGAGGATTCATCTGTACAAAGATCACTACATTTTCCAGGCAATTTCAAAATTTAT 1080
1021 GGTGAGGATTCATCTGTACAAAGATCACTACATTTTCCAGGCAATTTCAAAATTTAT 1080
1081 CTCCGTTATGGTCTGTAAGTAACTCGGTGAAAGGCGCACACCTCTGTCTTGAAGAAC 1140
1081 CTCCGTTATGGTCTGTAAGTAACTCGGTGAAAGGCGCACACCTCTGTCTTGAAGAAC 1140
1141 ACTGCACTGGAACATCTCGCTTGGATCTGTCTTCCATATATCAAGGCTTCTGAGTGA 1200
1141 ACTGCACTGGAACATCTCGCTTGGATCTGTCTTCCATATATCAAGGCTTCTGAGTGA 1200
1201 GCACACGATGAGTCCATTAAG 1224
1201 GCACACGATGAGTCCATTAAG 1224

RESULT 2
US-10-240-709-1
; Sequence 1, Application US/10240709
; Publication NO. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-10-240-709-1
Query Match 100.0%; Score 1224; DB 13; Length 1224;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCAGTCCCTGCTGATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTCTAAGA 60
DB 1 TCCAGTCCCTGCTGATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTCTAAGA 60
QY 61 ACAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB 61 ACAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
QY 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 181 ACTAGCAGGCAAGTGAAGAGCGGGCTGCTGAGTACCTGCTGAGTACCTGCTGAGTACCT 240
DB 181 ACTAGCAGGCAAGTGAAGAGCGGGCTGCTGAGTACCTGCTGAGTACCTGCTGAGTACCT 240
QY 241 ACCTGGAACAAGCCTCTACCTGATTAAACAGACAGGAGTCTCTGGGAGGACCTTA 300
DB 241 ACCTGGAACAAGCCTCTACCTGATTAAACAGACAGGAGTCTCTGGGAGGACCTTA 300
QY 301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATAATACAGAT 360
DB 301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATAATACAGAT 360
QY 361 GGTGTCCTGATCTGTGATTAACAGACAGATATGGACATGCCAAAGTCTGCTGTA 420
DB 361 GGTGTCCTGATCTGTGATTAACAGACAGATATGGACATGCCAAAGTCTGCTGTA 420
QY 421 TGGACACAGACGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480
DB 421 TGGACACAGACGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480
QY 481 GGCTATGAGTACCAACCTCGGGAGCGGATATGGACAAACCGGCTATCTGCACTTTC 540
DB 481 GGCTATGAGTACCAACCTCGGGAGCGGATATGGACAAACCGGCTATCTGCACTTTC 540
QY 541 GGTATGAGCAGTTTGAAGAAATAGATTTTTCAGTCTTGTGCTTACCAACCAATCTAC 600
DB 541 GGTATGAGCAGTTTGAAGAAATAGATTTTTCAGTCTTGTGCTTACCAACCAATCTAC 600
QY 601 AGCTGCTTCACTTCAAGCAGCTTCTGCTGAACTTATCTACATGCTCCCGGATGTGTC 660
DB 601 AGCTGCTTCACTTCAAGCAGCTTCTGCTGAACTTATCTACATGCTCCCGGATGTGTC 660
QY 661 AACTCCAGTTTCCCTTAAAGATCCCTGCTCGGTAACCTGCTGAACTGCTGAGGATGG 720

QY 31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTCTTTCTTTGCTCTCTTCTTCCCTC 90
Db 37 AAACAGAAAATGATGGCAAGACTGCTAAGAACATCCCTTGTCTTCTCTCTTCTTCCCTC 96
QY 91 TCTGGGCTCTGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
Db 97 TTTGGGCTCTGGGACACACAAATTTTCATGAGAAATGAAGAAAGGAAAGCTGTGGAC 156
QY 151 TGGTTATCTTTTATAGTTTACCAAGGACTAGCAAGGCAAGTGAAGAGGCGGGCTG 210
Db 157 TGGTTATCTTTTATAGTTTACCAAGGCAAGTGAAGAGGCGGGCTG 216
QY 211 CAGTACTGTACTGAGCTCCACAGAACCTGGAAACAGAGCCCTCTACCTGATTAAC 270
Db 217 GAGTACTGTACTGAGCTCCACAGAACCTGGAAAGAGTGAAGCACTAATGAAT 276
QY 271 AGCACCAGAGTGTCTCTGGGAGGACCTACAGCATCTGTATGACACACATATTCACG 330
Db 277 GACACCAAGAGTGTCTGGGAGGACATTAACACAGCTATATGAAGCATATGCTCTAAG 336
QY 331 AATG---ACACAGCCCTATCTAATATACACGATGTGTCTGATCTGTGAATTAACG 387
Db 337 AGTAACACACAGCTCTCTAATATACATGATGGAGTCCCTAACTGTGAATTAAGT 396
QY 388 AGCAGATATGGACATGCCAAAGTCTGTGGTATGAAACAGAGCGGGTCTGGCTG 447
Db 397 AGAAGATATGGACACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTCTGGCTG 456
QY 448 ATACACTCTGTCCCAAGTTTCCCGAGTTC-----ATGGCTATGAGTACCCACCTCG 501
Db 457 ATTCAATCCATCCCTCAGTTTCTCTCAATTCGGAGAGGCTATGATCCACCCACA 516
QY 502 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTCGGATACAGCCAGTTTGAGGAA 561
Db 517 GGGAGGCGATATGGACAAAGTGGCTCTGCAATACTTCAAGTACACAGGATGAGGCA 576
QY 562 ATAGATTTTACGCTCTTGGTCTTACACCAACATCTACAGCTGTCTTCAATTCAGACACC 621
Db 577 ATAGATTTTACGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCTTCAATTCAGACACC 636
QY 622 TTTCACTGGAACTTATCTACATGCCCGGATGTGTGCAACTCCAGTCTCTTAAAGATC 681
Db 637 TTTCACTGGAGCTCTATCTACATGCCCGGATGTGTGCAACTCCAGTCTCTTAAAGATC 696
QY 682 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCAGGCTCTAAACTCTGCTCAATTTGCA 741
Db 697 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCAGGCTCTAAACTCTGCTCAATTTGCA 756
QY 742 AAATCAAGTTTATCTGATGACATCTTTACAGGATGGATAGCTCAAAAGTTGAAGACA 801
Db 757 AAGTCGGATTTCTTTTTCAGGACATCTTTGAGGCTGGATGGCTCAACGGCTGAAGACA 816
QY 802 CATTTGTTAGCACAACCTGCGCAAGAAAGCAAGAGCTTCTTCAACTGTTCCCTG 861
Db 817 CACTTGTAAACAGAACTTGGCGGAAAGCAAGAGCTTCTTCAAACTGCTCCCTT 876
QY 862 CTTTCAATGTTCAACATCAAGTCCATTTGGGCTTAACTTCAAGTCTTCTTCAAGTCT 921
Db 877 CTTTCAATGTTCAACATCAAGTCCATTTGGGCTTAACTTCAAGTCTTCTTCAAGTCT 936
QY 922 CGCCAAGACCAATTCAAATGTTGTTTCCATATAAGGGCTCCGCAATTCGCTGACCTGC 981
Db 937 TATCAAGATCACGCAAGTGTGTATTTCCCAAGGGCACCAAAATTCGCTGACATGT 996
QY 982 ATTGGAGACCTTAAATCGAAGCTTACCAAGCCCTTAAGAGTGAAGATTCATCTGTACA 1041
Db 997 ATTGGAGACCTTAAATCGAAGCTTACCAAGCCCTTCAAGATGAGGATTCATTTGTACC 1056
QY 1042 AAGAATCACTACATTTACCGGCAATTTCAAAATTAATCTCCGTTATGGTCTGTAAAG 1101
Db 1057 CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATGATGATTAATCTATGAAGCTGTAA 1116
QY 1102 TAACTCGGTGAAGGCCACACCCCTCTGTCTTTGAAACACATGCACTGGCAATCTCGC 1161

Db 1117 TAACTTGGTGAAGGACACAGTACTATCATTTGAAAACTTGACATGGTCTTCTTCC 1176
QY 1162 CTTGGATCTCTTCTCCATAATA 1183
Db 1177 ATTACACCTCTTTATATTTTA 1198
RESULT 6
US-10-240-709-3
; Sequence 3, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-240-709-3
Query Match 55.0%; Score 672.8; DB 13; Length 1268;
Best Local Similarity 75.0%; Pred. No. 5.2e-208;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;
QY 31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTCTTTCTTTGCTCTCTTCTTCCCTC 90
Db 37 AAACAGAAAATGATGGCAAGACTGCTAAGAACATCCCTTGTCTTCTCTCTTCTTCCCTC 96
QY 91 TCTGGGCTCTGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
Db 97 TTTGGGCTCTGGGACACACAAATTTTCATGAGAAATGAAGAAAGGAAAGCTGTGGAC 156
QY 151 TGGTTATCTTTTATAGTTTACCAAGGACTAGCAAGGCAAGTGAAGAGGCGGGCTG 210
Db 157 TGGTTATCTTTTATAGTTTACCAAGGCAAGTGAAGAGGCAAGTGAAGAGGCTGTGGAC 216
QY 211 CAGTACTGTACTGAGCTCCACAGAACCTGGAAACAGAGCCCTCTACCTGATTAAC 270
Db 217 GAGTACTGTACTGAGCTCCACAGAACCTGGAAAGAGTGAAGCACTAATGAAT 276
QY 271 AGCACCAGAGTGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTCACG 330
Db 277 GACACCAAGAGTGTCTGGGAGGACATTAACACAGCTATATGAAGCATATGCTCTAAG 336
QY 331 AATG---ACACAGCCCTATCTAATATACACGATGTGTCTGATCTGTGAATTAAC 387
Db 337 AGTAACACACAGCTCTCTAATATACATGATGGAGTCCCTAACTGTGAATTAAGT 396
QY 388 AGCAGATATGGACATGCCAAAGTCTGTGGTATGAAACAGAGCGGGTCTGGCTG 447
Db 397 AGAAGATATGGACACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTCTGGCTG 456
QY 448 ATACACTCTGTCCCAAGTTTCCCGAGTTC-----ATGGCTATGAGTACCCACCTCG 501
Db 457 ATTCAATCCATCCCTCAGTTTCTCTCAATTCGGAGAGGCTATGATTCACCCACA 516
QY 502 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTCGGATACAGCCAGTTTGAGGAA 561
Db 517 GGGAGGCGATATGGACAAAGTGGCTCTGCAATACTTCAAGTACACAGGATGAGGCA 576
QY 562 ATAGATTTTACGCTCTTGGTCTTACACCAACATCTACAGCTGTCTTCAATTCAGACACC 621
Db 577 ATAGATTTTACGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCTTCAATTCAGACACC 636
QY 622 TTTCACTGGAACTTATCTACATGCCCGGATGTGTGCAACTCCAGTCTCTTAAAGATC 681
Db 637 TTTCACTGGAGCTCTATCTACATGCCCGGATGTGTGCAACTCCAGTCTCTTAAAGATC 696
QY 682 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCAGGCTCTAAACTCTGCTCAATTTGCA 741
Db 697 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCAGGCTCTAAACTCTGCTCAATTTGCA 756
QY 742 AAATCAAGTTTATCTGATGACATCTTTACAGGATGGATAGCTCAAAAGTTGAAGACA 801
Db 757 AAGTCGGATTTCTTTTTCAGGACATCTTTGAGGCTGGATGGCTCAACGGCTGAAGACA 816
QY 802 CATTTGTTAGCACAACCTGCGCAAGAAAGCAAGAGCTTCTTCAACTGTTCCCTG 861
Db 817 CACTTGTAAACAGAACTTGGCGGAAAGCAAGAGCTTCTTCAAACTGCTCCCTT 876
QY 862 CTTTCAATGTTCAACATCAAGTCCATTTGGGCTTAACTTCAAGTCTTCTTCAAGTCT 921
Db 877 CTTTCAATGTTCAACATCAAGTCCATTTGGGCTTAACTTCAAGTCTTCTTCAAGTCT 936
QY 922 CGCCAAGACCAATTCAAATGTTGTTTCCATATAAGGGCTCCGCAATTCGCTGACCTGC 981
Db 937 TATCAAGATCACGCAAGTGTGTATTTCCCAAGGGCACCAAAATTCGCTGACATGT 996
QY 982 ATTGGAGACCTTAAATCGAAGCTTACCAAGCCCTTAAGAGTGAAGATTCATCTGTACA 1041
Db 997 ATTGGAGACCTTAAATCGAAGCTTACCAAGCCCTTCAAGATGAGGATTCATTTGTACC 1056
QY 1042 AAGAATCACTACATTTACCGGCAATTTCAAAATTAATCTCCGTTATGGTCTGTAAAG 1101
Db 1057 CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATGATGATTAATCTATGAAGCTGTAA 1116
QY 1102 TAACTCGGTGAAGGCCACACCCCTCTGTCTTTGAAACACATGCACTGGCAATCTCGC 1161

577	ATAGATTCTCAGCTCTGGTCTGCAACCCCAACGCTTAGTGTCTCATCCACGACC	636
622	TTTCACTGGAAACTTATCTACATGCCGCCGATGTGTGCCAACTCCAGTTTCTTAAAGATC	681
637	TTTTCACAGAGAGCTCATCTACATGCCGCCGAGTGTGCACAGGGCCAGCTCATCAGAGATT	696
682	CCCTGTCGGGTACCTTCGCTGAACCTGCATCAGCCCGGGTCTTAACTTCGTCCATTTTGCA	741
697	CCCTGGCAGGCTCCTCACCACTTCACTCGGCCCCAGGCAAAAATTTCTCTCATTTTGA	756
742	AAATCAAGTTTTTATACCTGTATGATCATCTTTACAGGATGGATAGCTCAAAGTTTGAAGACA	801
757	AAGTCGGATTCTTTTCTTTGACGACATCTTTTCGAGCCTGATGTGCTCAA CGSCTGAAGACA	816
802	CATTGTGTAGCACAAACCTGGCGAAAAAGAAACAAGAGCTTCCTTCAAACCTTCCCTG	861
817	CACTTGTTAACAGAAACCTGCGACGAAAAAGACAAGAGCTTCCTTTCAAACCTGCTCCCTT	876
862	CCTTTACCATGTCTCAATCAAGTCAATTCATGGGGTAACTTCCAAAGTCTTACTTCAGTTCT	921
877	CCTTTACCATGTCTCAATATAAAAGCAATTAAATTTATCAGCACACTCTTATTTTCAGTTCT	936
922	CGCCAGACCACTTCCAAATGTGTGTTTCCATAAAGGGCTCCGCAAAATCGCTCGAACCTGC	981
937	TATCAAGATACGCCAAGTGTGTATTTTCCAAAGGGCCACCAAAATCGCTGGACATGT	996
982	ATTGGAGACCTTAATTCGAAGCCTACCAACGCCCTTAAAGGTGGAGGATTCATCTGTACA	1041
997	ATTGGAGACCTTAATTCGGAGTCCACACCAAGCCTTCAGAAATGGAGGATTCATTTGTACC	1056
1042	AAGAAATCACTACATTTACAGGCAATTCATAAATATATCTCCGTTATGGTTCTGTGAAG	1101
1057	CAGAAATGGCAAAATTCACCAAGCATTTCAAGGATTAGTATTTATCTAATGAAGAGCTGTAA	1116
1102	TAAACTCGGTGAAGGCCACACCCCTCTCTCTTGGAAAAACAATGGCACTGGAAACATCTCGC	1161
1117	TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAACCTTGTGACATGGGTCTTCTTCC	1176
1162	CTTGGATCTGTTCTCCATAATA	1183
1177	ATTACACCTTCTTTATATTTTA	1198

RESULT 7

```

RESULTS /
US-10-790-589-3
; Sequence 3, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-790-589-3
Query Match 55.0%; Score 672.8; DB 17; Length 1268;
Best Local Similarity 75.0%; Pred. No. 5.2e-208;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2
Ov 31 AGATAGAAATGACGACGAAAGCCTCTAAGAACAGTTCTTTCTTTGCTCTTCTTTGCCCTC 90

```

Db 1117 TAAACTCGTGAAGGACACAGGTACTATCAATTGAAAACCTTGACAATGGGTCTTCTTCC 117
 QY 1162 CTTGGATCTGTTCTCCATAATA 1183
 Db 1177 ATTACACCTTCTTTATATTTA 1198
 RESULT 8
 US-10-670-863-4
 ; Sequence 4, Application US/10670863
 ; Publication No. US20040157239A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tanuma, Sei-ichi
 ; APPLICANT: Shokawa, Daisuke
 ; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Th
 ; FILE REFERENCE: 224589
 ; CURRENT APPLICATION NUMBER: US/10/670,863
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: JP 11-230870
 ; PRIOR FILING DATE: 1999-08-17
 ; PRIOR APPLICATION NUMBER: US 09/807,784
 ; PRIOR FILING DATE: 2001-04-17
 ; NUMBER OF SEQ ID NOS: 15
 ; SEQ ID NO 4
 ; LENGTH: 1086
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1083)
 US-10-670-863-4
 Query Match 52.4%; Score 641.8; DB 17; Length 1086;
 Best Local Similarity 75.9%; Pred. No. 6.2e-198;
 Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;
 QY 31 AGATAGAAATGACGCAAGACCTCTAAGAACAGTCTTTCTTTGCTCTTCTTTGCCCTC 90
 Db 4 AACAAGAAATGATGGCAAGACTGTCTAAGAACATCTTTGCTTTGCTCTTCTTTGCCCTC 63
 QY 91 TCTGGGCTCTGGGGACACCAGAAATCTCATGACAGAAATGAATATGGTGAAGCTGTGAC 150
 Db 64 TTGGGGTCTGGGGGACGACACATTTTCATGAGAAATGAAGAAAGGGAAGCTGTGAC 123
 QY 151 TGGTTATCTTTTATAGTTTACCCAAAGGACTAGCAAGCAAGTGAAGAGCGGGGCTG 210
 Db 124 TGCTTTACTTTTATAAGTTTACCTAAAGACAAAACAAGAAATGGAGAGACTGGGTTA 183
 QY 211 CAGTACTGTACTGTGACCTCCACAGACAAAACCTGGAACAAGAGCTCTACTGATTAAAC 270
 Db 184 GAGTACTGTACTGTACTCTAACAAGCTGGAGAAAGTGAACAATAATGAAT 243
 QY 271 AGCACCAGAGTGCTCTGGGGAGGACCTTACAGCATCTGTATGACACACATAATTCACG 330
 Db 244 GACACCAAGAGTGTGTTGGGAGGACATTTACACAGCTATATGAAGCATATGCCCTCTAAG 303
 QY 331 ----ATGACACAGCCTTATCTAATATACACAGATGGTGTGCTTGGATCTGTGAATTACAGC 387
 Db 304 AGTAAACACACAGCCTATCTAATATACAAATGATGAGTCCCTTAACCTGTGAATTACAGC 363
 QY 388 AGACAGTATGGCATGCCAAGGCTGTGTGGTATGGAAACAGAACGAGGGTCTTGGCTG 447
 Db 364 AGAAGATATGGACACACCAAGGTTTACTGTGTGGAACAGAGTTCGAAGGGTCTTGGCTG 423
 QY 448 ATACACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGATACCCACCTCG 501
 Db 424 ATTCAATCCATCCCTCAGTTTCTCCAATTCCGGAAGAGGCTATGATATTCACCCACA 483
 QY 502 GGGAGCGATATGGACAAACCGGCATCTGCATCACTTTTCGGATACAGCCAGTTTGAGAA 561
 Db 484 GGGAGACGAAATGGACAAAGTGGCACTGCACTAATTTCAAGTACACCAAGTATGAGCA 543
 QY 562 ATAGATTTTCAGCTCTTGGTCTTACCAACAACATCTACAGCTGCTTATTTCCAGACCC 621

Fri Oct 15 10:01:01 2004

US-10-790-589-1.rnpb

391 CAGTATGACATGCGAAGGCTGCTGATGGAACAGACGAGGGTCTGGCTGATA 450
231 AAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTCTGGCTGAT 290
451 CACTCTGTTCCCAAGTTTCCCGAGTTC-----ATGGCTATGAGTACCCAACTCGGGG 504
291 CATTCATCCCTCAGTTTCTCCAAATTCGGGAAGAGCTATGATTATCCACCCACAGG 350
505 AGCGGATATGGAACAAACCGGCTCTGATCATCTTCGGATACAGCCAGTTTGAAGAAATA 564
351 AGACGAATGGAACAAAGTGGCATCTGCATAAACTTTCAAGTACAACAGATGAGGCAATA 410
565 GATTTTCAGCTCTGTTCTTACAAACAACTACAGCTGCTTCAATCCAAAGCACCTTT 624
411 GATTCAGCTCTGTTCTGTCGACACCCCAAGCTATATAGCTTGTCTCCATCCAGGACCTTT 470
625 CACTGGAA 632
471 CACCAGGA 478
RESULT 10
US-09-908-975-28714
; Sequence 28714, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28714
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-28714
Query Match 5.3%; Score 65; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.7e-10; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 0
QY 1118 CCACACCTCTGCTCTTGAACAACTGGCACTGGAACATCTCGGCTTGGATCTGTTCTCC 1177
Db 1 CCACACCTCTGCTCTTGAACAACTGGCACTGGAACATCTCGGCTTGGATCTGTTCTCC 60
QY 1178 ATAAT 1182
Db 61 ATAAT 65
RESULT 11
US-10-408-167A-1
; Sequence 1, Application US/10408167A
; Publication No. US20030219428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way

CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/408,167A
FILING DATE: 04-Apr-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/861034
FILING DATE: 18-May-2001
APPLICATION NUMBER: 08/639294
FILING DATE: 25-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Evans, David W.
REGISTRATION NUMBER: NONE
REFERENCE/DOCKET NUMBER: P1024D1C1
TELEPHONE: 650/225-1739
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-408-167A-1
Query Match 5.0%; Score 60.8; DB 16; Length 1575;
Best Local Similarity 44.8%; Pred. No. 3.2e-08;
Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;
QY 200 AGCGGGGCTGCAGTACCTGTACCTGGACTCCCAAGACAAACCTGGAAACAGAGCTCT 259
Db 222 AGAGAGGGCTGCAGTACAGTATCTGACGAGAGCTCCGGAGGCTGGCGGAGCGGAGG 281
QY 260 ACCTGATTAACAGACAGAGAGTCTCTGGGAGGAGCTTACAGCATCTGTATGACACAC 319
Db 282 CACTCATCAACAGCCCGAGGGGCGCTGGGCGGAGCCCTGCAGCCGCTGTACCGGAGCA 341
QY 320 ATATTCACGAATGACACAGCCCTATCTAATATACAAAGATGGTGTCCCTGGATCTGTGA 379
Db 342 ACACCGAGGCTGCGCTTCTCTGCTCTACATGACCAACCGCCTCAACCCAGCAGGCTC 401
QY 380 ATTACAGCAGACAGTATGGACATGCGCAAGAGTCTGCTGGTATGGAACAGAACGAGGGT 439
Db 402 AGGACTCTTCATCGTGGGACACGAAAGGTGTCTCTGCTCCTTGACCAACGATGGGGCT 461
QY 440 TCTGGCTGATACATCTGTTCCCAAGTTTCCCGAGTTCATGGCTATGAGTACCCAGCT 499
Db 462 TCTGGCTGTTCCACAGTGTACCTAACTTCCCTCCACCGGCGCTCTCTGTGCTGATACAGCT 521
QY 500 CG-----GGGAGGCGATATGGAACAAACCGGCACTCTGCATCACTTCCGATACAGCC 550
Db 522 GGCCTCATVAGCGCTGTACCTACGCGGAGAGCCCTGCTCTGTGTGCTCTTCCCTTGGCTC 581
QY 551 AGTTTGAGGAATAAGATTTTTCAGCTCTTGGTCTTGTACCAACCAACATCTACAGCTGTCTCA 610
Db 582 AGTTCTCGAAGATGGCAAGCAGCTGACCTACACCTACCCCTGGGTCTTATAACTACCCAGC 641
QY 611 TTCCAGCACCTTTTCACTGGAACTTATCTACATGCGCCCGGATGTGTGCCAACTCCAGTT 670
Db 642 TGGAGGGGATCTTTGCCCAGGAATTTCCCGAGCTTGGAGAATGTGGTCAAGGGCCACACAG 701
QY 671 CCTTAAAGATCCCTGTCCGGTACCTCTGCTGAGCTGCACTACAGCCAGGGTCTTAAACTCG 730

Db	702	TTAGCCAAAGAACCCCTGGAAACAGCAGCATCACACTCACATCCAGCGCGGGCTGTTTTC	761
QY	731	TCCATTTTGC AAAATCAAGTTTTTATCTACTGATGATCATCTTTACAGGATGGA TAGCTCAA	790
Db	762	AGAGCTTTGCCAAGTTCTCAGCAAATTTGGAGATGACCTGTACTCCGGCTGTGTTGCAGCAG	821
QY	791	AGTTGAAGACACATTTGTTAGCACAAACCTGGCGAAAAAACAAGAGTTCCTTCAA	850
Db	822	CCCTTGGTACCAACCTGCAGGTCAGATTCTGGCACAAAACGTGAGGCATCTGCGCCTCTA	881
QY	851	ACTGTTCCCTGCCCTTACCATGTCTACAACTCAAGTCAATGGGGT-----AACTTCCA	904
Db	882	ACTGTCGGATATCTGGCAGGTTCTGAAATGTGAACCAAGTAGCTTTCCTGGAACGACCG	941
QY	905	AGTCTTACTTTCAGTTCTCGCCAAAGACCATTCCAAATGGTGTGTTTCCATAAAGGCTCCG	964
Db	942	GCCCAAGTTTCAAGCAGCAGAGGACCATTCCAAATGGTGGTGTCCCTCCCAAAAGG-----	996
QY	965	CAAAATCGCTGGACCTGCATTGGAGACCTTAAATCGAAGCCTACACCAAGCCTTAAGAGGTG	1024
Db	997	----GCCCTGGACCTGCTGGTGACATGAATCGGAACGAGGAGAGGACCAACGGGTG	1052
QY	1025	GAGGATTCATCTGTACAAAGATCACTACTACATTACAGGCAATTTC	1070
Db	1053	GGGCAACACTGTGTGCCAGCTGCAGCCCTCTGGAAAGCTTCCA	1098

```

RESULT 12
US-09-908-975-2418
; Sequence 2418, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2418
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-908-975-2418

```

	Query Match	4.4%;	Score 53.8;	DB 10;	Length 65;
	Best Local Similarity	89.2%;	Pred. No. 7.9e-07;		
Matches	58; Conservative	0; Mismatches	7; Indels	0; Gaps	0;
QY	969	TGCTGGACCTGATTGGAGACCTTAATCGAAGCCTACACCAAGCCTTTAGAGGTGGAGG	1028		
Db	1	TCGCTGGACTGTATTGGAGACCTTAATCGGAGCCCCACATCAAGCCTTGAGAATGGAGG	60		
QY	1029	ATTCA	1033		
Db	61	ATTTA	65		

RESULT 13
US-09-908-975-13222
; Sequence 13222, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi

```

: APPLICANT: WASERMAN, Alon
: APPLICANT: MINTZ, Eli
: APPLICANT: MINTZ, Eli
: APPLICANT: FAIGLER, Simchon
: TITLE OF INVENTION: OLIGONUCLEOTIDE
: TITLE OF INVENTION: THAT POPU
: FILE REFERENCE: 36688-0005
: CURRENT APPLICATION NUMBER: US
: CURRENT FILING DATE: 2001-07-2
: PRIOR APPLICATION NUMBER: US 6
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 6
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 32337
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 13222
: LENGTH: 60
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-098-975-13222

```

	Query March	4.0%	Score 48.8;	DB 10;	Length 60;
	Best Local Similarity	88.3%;	Pred. No. 3.2e-05;		
	Matches 5;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	971	GTGGGACCTGCATTTGGAGACCTAAATCGAAGCCCTACACCAAGCCTTAAGAGGTGGAGGAT	1030		
Db	1	GTGGGACATGTA TTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAGAAAGTTGGAGGAT	60		

```

RESULT 14
US-09-960-352-7234
; Sequence 7234, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wyatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7234
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 31-LIB3058-051-Q1-K1-H3
US-09-960-352-7234

```

Query Match	3.3%;	Score 41;	DB 9;	Length 379;
Best Local Similarity	55.1%;	Pred. No. 0.037;		
Matches 102;	Conservative 0;	Mismatches 80;	Indels 3;	Gaps 1;
QY	130	GAATATGGTGAAGCTGTGACCTGGTTTACTTTTATAGTTACCCAAAGGACTAGCAAG	189	
Db	118	GACTCGGCGCAGCTGTGACCTGGTTCTGCGATTACAAGCTGCCGGCCACACAGGGTCG	177	
QY	190	GCAAGTG---AAGAGCGGGGTGCACTAGTACCTGTACTCGACTTCCAAGACACAAACCTGG	246	
Db	178	GGAGATGCGACGCAGAAACGGTCTGCGGTCAAGTACTTTGACGAACACTCATAGACTGG	237	
QY	247	AACAAGAGCCTTACTCTGATTAAACAGCACAGAGGTGCTCTGGGAGGACCTTACAGCAT	306	
Db	238	ACGACGCGGTGGGGTCTATCATATGACACACACGGGTGCCGTGGGCCGCAGACTGCTGCCG	297	
QY	307	CTGTA 311		
Db	298	-CTGTA 302		

Fri Oct 15 10:01:01 2004

RESULT 15
US-10-311-455-1276/c
; Sequence 1276, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1276
; LENGTH: 6609
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1276

Query Match 3.2%; Score 39; DB 15; Length 6609;
Best Local Similarity 58.0%; Pred. No. 1;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 733 CATTTTGCAAAATCAAGTTTTTATCTGATGACATCTTTACAGGATGGATAGCTCAAAG 792
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5914 CATTTTAAATCTAATTATATATATAAAAAAACCAAAAAAACTCTAATATATACCAAT 5855

Qy 793 TTGAAGACATTTGTTAGCACAACTGGCAGAAAAGAAACAAAGAGCTTCCTTCAA 851
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5854 TTACAAAAATATTATTATTAACATTCCTGATATAAAACAAAAATTACAAAACCTCCAA 5796

Search completed: October 14, 2004, 08:18:15
Job time : 597.827 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:43:42 ; Search time 103.146 Seconds
(without alignments)
6585.417 Million cell updates/sec

Title: US-10-790-589-1

Perfect score: 124

Sequence: 1 tccagtcctccctgcattgaa.....aacgtagctccaataaaag 1224

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUTS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1224	100.0	1224	4	US-09-574-942-1
2	1219.8	99.7	1652	4	US-09-807-784B-2
3	672.8	55.0	1268	4	US-09-574-942-3
4	641.8	52.4	1086	4	US-09-807-784B-4
5	60.8	5.0	1575	3	US-08-639-294-1
6	60.8	5.0	1575	4	US-09-861-034B-1
7	60.8	5.0	1915	3	US-09-147-915-1
8	55.4	4.5	927	3	US-09-147-915-2
9	36.2	3.0	1230025	4	US-09-198-452A-1
10	36	2.9	7218	1	US-08-232-463-14
11	35.4	2.9	1005	4	US-09-547-435-7
12	35.4	2.9	1383	4	US-09-547-435-11
13	35.4	2.9	1470	4	US-09-547-435-3
14	35.4	2.9	1848	4	US-09-547-435-9
15	35.4	2.9	1938	4	US-09-547-435-25
16	35.4	2.9	2236	4	US-09-547-435-5
17	35.4	2.9	2316	4	US-09-547-435-27
18	35.4	2.9	2604	4	US-09-547-435-23
19	35.4	2.9	2701	4	US-09-547-435-1
20	35.4	2.9	3384	4	US-09-547-435-29
21	35	2.9	1230025	4	US-09-198-452A-1
22	34.8	2.8	262	4	US-09-313-294A-2280
23	33.8	2.8	640681	4	US-09-790-988-1
24	33.4	2.7	1828	4	US-09-744-900-2
25	33.4	2.7	3742	4	US-09-744-900-1
26	33.2	2.7	396	4	US-09-640-173-140
27	33.2	2.7	396	4	US-09-713-550-140

28	33.2	2.7	10432	4	US-09-919-172-97	Sequence 97, Appl
29	33.2	2.7	10432	4	US-09-976-594-21	Sequence 21, Appl
30	33	2.7	3321	1	US-08-484-438-5	Sequence 5, Appl
31	33	2.7	5484	3	US-09-632-580A-3	Sequence 1, Appl
32	33	2.7	5501	1	US-08-484-438-1	Sequence 1, Appl
33	32.6	2.7	1773	4	US-09-495-050A-51	Sequence 51, Appl
34	32.4	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl
35	32.2	2.6	1582	4	US-09-210-422-4	Sequence 14, Appl
36	32	2.6	396	4	US-09-640-173-56	Sequence 56, Appl
37	32	2.6	396	4	US-09-713-550-56	Sequence 56, Appl
38	32	2.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
39	32	2.6	2196	3	US-09-142-759-2	Sequence 2, Appl
40	32	2.6	2258	4	US-09-976-594-1023	Sequence 1023, Ap
41	31.8	2.6	445	4	US-09-621-976-14778	Sequence 14778, A
42	31.8	2.6	531	4	US-09-252-991A-12110	Sequence 12110, A
43	31.8	2.6	1011	4	US-09-252-991A-12311	Sequence 12311, A
44	31.8	2.6	1539	4	US-09-489-039A-2429	Sequence 2429, Ap
45	31.6	2.6	312	4	US-09-252-991A-12407	Sequence 12407, A

ALIGNMENTS

RESULT 1

US-09-574-942-1

; Sequence 1, Application US/09574942

; Patent No. 6358723

; GENERAL INFORMATION:

; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/574,942

; CURRENT FILING DATE: 2000-05-19

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1224

; TYPE: DNA

; ORGANISM: Mus sp.

US-09-574-942-1

Query Match Similarity 100.0%; Score 1224; DB 4; Length 1224;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCCAGTCCCTGCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTAAGA	60
Db	1	TCCAGTCCCTGCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTAAGA	60
Qy	61	ACAGTCTCTTTTGTCTCTTTTGTCCCTCTCTGGGGTCTCTGGGGACACCAAGAAATCTCA	120
Db	61	ACAGTCTCTTTTGTCTCTTTTGTCCCTCTCTGGGGTCTCTGGGGACACCAAGAAATCTCA	120
Qy	121	TGCAGAAATGAATATGCTGAAGCTGTGGAATCTTTTATTAAGTTACCAAAAGG	180
Db	121	TGCAGAAATGAATATGCTGAAGCTGTGGAATCTTTTATTAAGTTACCAAAAGG	180
Qy	181	ACTAGCAAGCAAGTGAAGAGGGGGCTCAGTACCTGTGAGTCTCCAGACAA	240
Db	181	ACTAGCAAGCAAGTGAAGAGGGGGCTCAGTACCTGTGAGTCTCCAGACAA	240
Qy	241	ACCTGGAACCAAGACCTCTACCTGATTAACAGACACAGGAGTGTCTGGGAGACCTTA	300
Db	241	ACCTGGAACCAAGACCTCTACCTGATTAACAGACACAGGAGTGTCTGGGAGACCTTA	300
Qy	301	CAGCATCTGTATGACACACATAATTCACGAATGACACAGCCTATCTAATATACACGAT	360
Db	301	CAGCATCTGTATGACACACATAATTCACGAATGACACAGCCTATCTAATATACACGAT	360
Qy	361	GGGTCTCTGTGATCTGTGAATTACAGCAGACAGTATGGACATGCCAAAGTCTGCTGGTA	420

RESULT 4
US-09-807-784B-4
; Sequence 4, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use The
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15

QY	922	CGCCAAAGACCAATTC	CAAAATGGTGTGTTTCCATAAAGGGCTCCGCAAAATCGCTGGACCTGC	981
DB	904	TATCAAGATCATG	CCCAAGTGGTGTATTTCCCAAAAGGGGCACCAAAAATCGCTGGACATGT	963
QY	982	ATTGAGACCTAAATC	GAAAGCCTACACCAAGCCTTAAGAGGTGAGAGATTCACTCTGACA	1041
DB	964	ATTGAGACCTAAATC	GGAGTCCACACCAAGCCTTCAGAAATGGAGGATTCAATTTGATCC	1023
QY	1042	AGAGTCACTACATTT	TACCAGGCAATTCATAAATTATATCTCCGTTATGGGTTCTGTGAAG	1101
DB	1024	CAGAATTGGCAAAT	TATCCCAAGCAATTCAGGATTAATATATCTATGAAAGCTGTGAAG	1083
QY	1102	TAA	1104	
DB	1084	TAA	1086	
RESULT 5				
US-08-639-294-1				
; Sequence 1, Application US/08639294				
; Patent No. 6265195				
; GENERAL INFORMATION:				
; APPLICANT: Baker, Kevin P.				
; APPLICANT: Baron, Will F.				
; TITLE OF INVENTION: Human DNase II				
; NUMBER OF SEQUENCES: 9				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Genentech, Inc.				
; STREET: 460 Point San Bruno Blvd				
; CITY: South San Francisco				
; STATE: California				
; COUNTRY: USA				
; ZIP: 94080				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: WinPatIn (Genentech)				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/639,294				
; FILING DATE: 25-Apr-1996				
; CLASSIFICATION: 435				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Johnston, Sean A.				
; REGISTRATION NUMBER: 35,910				
; REFERENCE/DOCKET NUMBER: P1024				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 415/225-3562				
; TELEFAX: 415/952-9881				
; TELEX: 910/371-7168				
; INFORMATION FOR SEQ ID NO: 1:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1575 base pairs				
; TYPE: Nucleic Acid				
; STRANDEDNESS: Single				
; TOPOLOGY: Linear				
; MOLECULE TYPE: DNA				
US-08-639-294-1				
Query Match				
Best Local Similarity 5.0%; Score 60.8; DB 3; Length 1575;				
Matches 399; Conservative 0; Mismatches 467; Indels 24; Gaps 3				
QY	200	AGCGGGCGCTGCAGTACTGTGTA	CTTGACCTCCACAGCAACCTTGGACACAGAGCCCTCT	259
DB	222	AGAGGGCTGCAGTACAAAGTAT	CTGGACAGAGCTCCGGAGGCTGGCGGACCGCAGG	281
QY	260	ACCTGATTACAGCACCAAGGAGTCT	CTGGGGAGGACCTTACAGCATCTGTATGACACAC	319
DB	282	CACATCATCAACAGCCCGAGGGGCG	CGTGGCCGAGAGCTTCGACGCGCTGTACCGAGCA	341
QY	320	ATAATTCACGAATGACACAGCCTT	CTTAATATATACACGATGGTTCCTCGATCTGTGA	379

Db 882 ACTGCTGGATATCTGGCAGGTTCTGAATGTGAACCCAGATAGCTTTCCCTGGACACGCG 941
 QY 905 AGTCTTACTTCACTTCTGCGCAAGACCATTCCTAAATGGTGTGTTTCCATAAAGGGCTCCG 964
 Db 942 GCGCAAGCTTCAACAGCAGACAGAGACCACTCCAAATGGTGGTGTCTCCCAAAAGG---- 996
 QY 965 CAAATCGCTGGACCTCATTGGAGACCTAAATCGAAGCCTTACACCAAGCCTTAAGAGGTG 1024
 Db 997 ----GCCCTGGACCTCGCTGGTGTGATGATCGGAACACAGGAGAGGACCAAGCGGTG 1052
 QY 1025 GAGGATTCATCTGTACAAAGAAATCACTTACATTTACCAGGCAATTTCA 1070
 Db 1053 GGGGCACACTGTGTGCCAGCTGCCAGCCCTCTGGAAAGCCTTTCCA 1098

RESULT 7
 US-09-147-915-1
 ; Sequence 1, Application US/09147915A
 ; Patent No. 6184034
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan
 ; APPLICANT: Krieser, Ronald
 ; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
 ; FILE REFERENCE: DC-0097
 ; CURRENT APPLICATION NUMBER: US/09/147,915A
 ; CURRENT FILING DATE: 1999-03-23
 ; EARLIER APPLICATION NUMBER: PCT/US97/18262
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/028,539
 ; EARLIER FILING DATE: 1996-10-15
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1915
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-147-915-1

Query Match 5.0%; Score 60.8; DB 3; Length 1915;
 Best Local Similarity 44.6%; Pred. No. 4.8e-10;
 Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;

QY 200 AGCGGGGCTGACGTACTCTGCTGACTCCACAGACAAACCTTGAACAAAGAGCCTCT 259
 Db 204 AGAGAGGCTCGACTCAAGTAATCTGACGAGAGCTCGGAGGCTGGGGAGCGCAGG 263
 QY 260 ACTGATTAACAGCACAGAGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACAC 319
 Db 264 CACTCATCAACAGCCCGGAGGGGGCGTGGGGCCGAAAGCTGACGCGGTGTACCGGAC 323
 QY 320 ATAAATCCAGATGACACAGCCCTATCTAATATACAGATGGTGTCCCTGGATCTGTGA 379
 Db 324 ACACAGCCAGCTGCGCTTCTGTCTTACATGATGACCAACCGCCCTCAACCCAGAGGCTC 383
 QY 380 ATTACAGCAGACATGAGCATGCAATGCCAAAGTCTGCTGGTATGGAACAGAACGAGGGT 439
 Db 384 AGGACTCTTCATCGTGGGACACAGAGGCTGCTCTCTCTGACCAAGTGGGGCT 443
 QY 440 TCTGGCTGATACACTCTGTTTCCCAAGTTTCCCCAGT-----TCATGGCTATGAT 490
 Db 444 TCTGGCTGGTCCACAGTGTACTAACTTCCCTCCACCGGCTCTCTGCTGCATACAGTT 503
 QY 491 ACCCAACCTCGGAGGGGATATGACAAACCGGCATCTGCATCATCTTCGGATACAGCC 550
 Db 504 GGCCTCATAGCGCCTGTACTACCGGGCAGACCTGCTCTGTGTGTCTTTTCCCTTCGCTC 563
 QY 551 AGTTTGAAGAAATAGATTTTTCAGTCTTGTGTTTCTTACCAACCAATCTTACAGCTGCTTCA 610
 Db 564 AGTTCTGAAGATGGCAAGCAGCTGACCTACACCTACCTCCCTGGGTCTATAACTACGAC 623
 QY 611 TTCCAGCAGCTTTCACCTGGAAACTTATCTATACATGCCCGGATGTGTGCCAACTCCAGTT 670
 Db 624 TGAAGGGATCTTTTCCCGAGGAATTCCTCCGACTTGGAGAAATGTGGTCAAGGGCCACG 683

QY 671 CCTTAAAGATCCCTGTCTCGGTACTCTCGTGAATGCACTCAGCCCAGGCTTAAACTTCG 730
 Db 684 TTAGCCAAAGAACCTCGGAACAGCAGCATCACTCATATCCAGGCGGGCTGTTTTCC 743
 QY 731 TCAATTTGCAAAATCAAGTTTATCTATGATGATCTTTTACAGATGATAGCTCAAA 790
 Db 744 AGAGCTTTTCCCAAGTTTTCAGCAAAATTTGGAGATGACCTGTACTCCGGCTGGTGGCAG 803
 QY 791 AGTTGAAGACATATTTTGTAGCAAAACCTGGCAGAAAGAAACAAAGAGCTTCTTTCAA 850
 Db 804 CCCTTGGTACCAACCTGCGAGTCCAGTTCTGGCAAAACTGTAGCATCTCTGCCCTCTA 863
 QY 851 ACTGTTCCCTGCTTACCATGTCTCAACAATCAAGTCCATTTGGGGT-----AATTTCCA 904
 Db 864 ACTGCTCGGATATCTGCGAGTTCTGAATGTGAACCCAGATAGCTTTCCCTGGACCAAGCG 923
 QY 905 AGTCTTACTTCACTTCTCGCAAGACATTCCTCAATGGTGTGTTTCCATAAAGGGCTCCG 964
 Db 924 GCCCAAGCTTCAACAGCAGCAGAGGACCTCCAAATGGTGGTGTCTCCCAAAAGG----- 978
 QY 965 CAAATCGCTGGACCTGATTGGAGACCTAAATCGAAGCCTTACACCAAGCCTTAAGAGGTG 1024
 Db 979 ----GCCCTGGACCTGCTGGCCAGCTCCAGCTCCAGCCCTCTGGAAAGCCTTTCCA 1080
 QY 1025 GAGGATTCATCTGTACAAAGAAATCACTTACATTTACCAGGCAATTTCA 1070
 Db 1035 GGGGCACACTGTGTGCCAGCTCCAGCTCCAGCCCTCTGGAAAGCCTTTCCA 1080

RESULT 8
 US-09-147-915-2
 ; Sequence 2, Application US/09147915A
 ; Patent No. 6184034
 ; GENERAL INFORMATION:
 ; APPLICANT: Eastman, Alan
 ; APPLICANT: Krieser, Ronald
 ; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
 ; FILE REFERENCE: DC-0097
 ; CURRENT APPLICATION NUMBER: US/09/147,915A
 ; CURRENT FILING DATE: 1999-03-23
 ; EARLIER APPLICATION NUMBER: PCT/US97/18262
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/028,539
 ; EARLIER FILING DATE: 1996-10-15
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 927
 ; TYPE: DNA
 ; ORGANISM: Bos sp.
 US-09-147-915-2

Query Match 4.5%; Score 55.4; DB 3; Length 927;
 Best Local Similarity 46.3%; Pred. No. 2.1e-08;
 Matches 311; Conservative 0; Mismatches 336; Indels 24; Gaps 3;

QY 396 TGGCATGCCAAAGGCTGCTGTGATGGAACAGAACGAGGGGTTCTGGCTGATACACTC 455
 Db 75 TGGGCACACAGAGGCTGTGCTCTCTGGACCAAGAGGGGGCTTCTGGTTGATCCACAG 134
 QY 456 TGTTTCCCAAGTTTCCCCCAGTTTCAAGGTATGAGTACCCACCTCG-----GGGAG 506
 Db 135 CGTTTCCAAACTTCCCTCCACGCTGCTCTCTGCTGCGGTACAGCTGGGCTCTCTGGTCCCA 194
 QY 507 GCGATATGGAACAAACCGGCATCTGCATCACTTTCGGATACAGCCAGTTTGGAGAAATAGA 566
 Db 195 AAAATATGGGAGAGCCCTGATCTGTATCTTTCTCTCCACCAAGTTTCTGGATACAG 254
 QY 567 TTTTTCAGCTCTTGGTCTTTTCAACCAACCAATCTACAGCTGTCTTCAATTCACAGCACTTTCA 626
 Db 255 CAAACAGCTGACTTATACCTATCTGATATGATGACCAAGGCTGGAAAGGGGACTTTGG 314

TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 2.9%; Score 36; DB 1; Length 7218;
Best Local Similarity 8.2%; Pred. No. 0.43;
Matches 33; Conservative 188; Mismatches 183; Indels 0; Gaps 0;

QY 35 AGAAATGACAGCAAGCCTTAAGACAGTCTTCTTTCTCTCTTTCTTTGCGCCCTCTCTG 94
Db 1507 AAAAAAGGCGATGATGACATCACTGTAAATACCTATCTATGCAAGTAGTTAAAGAGATAG 1448
QY 95 GGGTCTCTGGGACACACAGAAATCTCATGACAGAAATGAATATGATGAGCTGTGACCTGGT 154
Db 1447 AAGAAATTGTACERRR 1388
QY 155 TTATCTTTTATAAGTTACCCAAAGGACTAGCAAGGCAAGTGAAGCGCGGCGCTGCACT 214
Db 1387 RR 1328
QY 215 ACCTGTACCTGGACTCCACAGACAAACCTGGACAGACCTCTACCTGATTAACAGCA 274
Db 1327 RR 1268
QY 275 CCAGGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATAATTCACGAATG 334
Db 1267 RR 1208
QY 335 ACACAGCCTATCTAATACAAACGATGCTGCTCTGATCTGTGAATTAACAGACAGACT 394
Db 1207 RR 1148
QY 395 ATGACATGCCAAGTCTGCTGTATGATGAACAGACGAGGGG 438
Db 1147 RR 1104

RESULT 11

US-09-547-435-7/c
Sequence 7, Application US/09547435
Patent No. 6582957
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander, Jr.
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
FILE REFERENCE: 7705.0009-00000
CURRENT APPLICATION NUMBER: US/09/547,435
CURRENT FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 1005
TYPE: DNA
ORGANISM: Homo sapiens
US-09-547-435-7

Query Match 2.9%; Score 35.4; DB 4; Length 1005;
Best Local Similarity 53.2%; Pred. No. 0.17;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 629 GGAACCTTATCTACATGCCCCGGATGTGTGCCAACTCCAGTTCCTTAAAGATCCCTGTCC 688
Db 166 GAACATTGGCTCCATGTATCATCAGGATGGATGTCAATTTTCATGGGAGCCGGCA 107
QY 689 GGTACTCGTGAATGCACTAGCCAGGGTCTAACTTCGTCATTTTGCAAAATCAA 748
Db 106 GGTACCGAATCCCACTGTCTGCTGTCTACACAAGTTGTCTCTTTGTCAAGGCAA 47

QY 749 GTTTTATATCTGATGACATCT 769
Db 46 ATTTCTGTCTGACTCCATCT 26

RESULT 12

US-09-547-435-11/c
Sequence 11, Application US/09547435
Patent No. 6582957
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander, Jr.
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
FILE REFERENCE: 7705.0009-00000
CURRENT APPLICATION NUMBER: US/09/547,435
CURRENT FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
US-09-547-435-11

Query Match 2.9%; Score 35.4; DB 4; Length 1383;
Best Local Similarity 53.2%; Pred. No. 0.21;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 629 GGAACCTTATCTACATGCCCCGGATGTGTGCCAACTCCAGTTCCTTAAAGATCCCTGTCC 688
Db 166 GAACATTGGCTCCATGTATCATCAGGATGGATGTCAATTTTCATGGGAGCCGGCA 107
QY 689 GGTACTCGTGAATGCACTAGCCAGGGTCTAACTTCGTCATTTTGCAAAATCAA 748
Db 106 GGTACCGAATCCCACTGTCTGCTGTCTACACAAGTTGTCTCTTTGTCAAGGCAA 47
QY 749 GTTTTATATCTGATGACATCT 769
Db 46 ATTTCTGTCTGACTCCATCT 26

RESULT 13

US-09-547-435-3/c
Sequence 3, Application US/09547435
Patent No. 6582957
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander, Jr.
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
FILE REFERENCE: 7705.0009-00000
CURRENT APPLICATION NUMBER: US/09/547,435
CURRENT FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1470
TYPE: DNA
ORGANISM: Homo sapiens
US-09-547-435-3

Query Match 2.9%; Score 35.4; DB 4; Length 1470;
Best Local Similarity 53.2%; Pred. No. 0.22;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

The invention relates to deoxyribonuclease (DNase) II beta proteins and their corresponding cDNAs. The DNase II beta may be useful to digest DNA in the mucous plugs in lungs of cystic fibrosis patients and so reduce their viscosity. The present sequence is mouse DNase II beta cDNA

```

XX SQ Sequence 1224 BP; 351 A; 308 C; 256 G; 309 T; 0 U; 0 Other;
Query Match 100.0%; Score 1224; DB 5; Length 1224;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAGTCCCTGATGGAATGAAGGCCACAGATAGAAATGACAGCAAGCCCTTAAGA 60
DB 1 TCCAGTCCCTGATGGAATGAAGGCCACAGATAGAAATGACAGCAAGCCCTTAAGA 60
QY 61 ACAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB 61 ACAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
QY 121 TGCAGAAATGAATATGATGAGCTGTGAGCTGTGTTATCTTTATTAAGTTACCAAGG 180
DB 121 TGCAGAAATGAATATGATGAGCTGTGAGCTGTGTTATCTTTATTAAGTTACCAAGG 180
QY 181 ACTAGCAAGGCAAGTGAAGAGCGGGGCTGAGTACCTGTGACTCCACAGACAA 240
DB 181 ACTAGCAAGGCAAGTGAAGAGCGGGGCTGAGTACCTGTGACTCCACAGACAA 240
QY 241 ACCTGGAACCAAGCCCTTACCTGATTAACAGCACAGGAGTGTCTTGGGAGGACCTTA 300
DB 241 ACCTGGAACCAAGCCCTTACCTGATTAACAGCACAGGAGTGTCTTGGGAGGACCTTA 300
QY 301 CAGCATCTGATGACACATAATTCACGAAATGACACAGCCTATCTTAATATACAGAT 360
DB 301 CAGCATCTGATGACACATAATTCACGAAATGACACAGCCTATCTTAATATACAGAT 360
QY 361 GGTGTCCCTGGATCTGTGAAATACAGCAGACAGTATGAGATGCAAGGTTCTCTGGTA 420
DB 361 GGTGTCCCTGGATCTGTGAAATACAGCAGACAGTATGAGATGCAAGGTTCTCTGGTA 420
QY 421 TGGAAACAGACGCGAGGGGTTCTGCTGTGATACCTCTGTTCCTCAAGTTTCCCCAGTTTCAT 480
DB 421 TGGAAACAGACGCGAGGGGTTCTGCTGTGATACCTCTGTTCCTCAAGTTTCCCCAGTTTCAT 480
QY 481 GGCTATGATGACCAACCTCGGGAGGCGATATGCAACACCGGATCTGATCATTTC 540
DB 481 GGCTATGATGACCAACCTCGGGAGGCGATATGCAACACCGGATCTGATCATTTC 540
QY 541 GGATACAGCCAGTTTGAAGAAATAGATTTTCAGCTCTTGTCTTACACCAACATCTAC 600
DB 541 GGATACAGCCAGTTTGAAGAAATAGATTTTCAGCTCTTGTCTTACACCAACATCTAC 600
QY 601 AGCTGCTTCATTCCCAAGCACCTTTTCACTGGAACCTTATCTACATGCCCGGATGTGCC 660
DB 601 AGCTGCTTCATTCCCAAGCACCTTTTCACTGGAACCTTATCTACATGCCCGGATGTGCC 660
QY 661 AACTCCAGTTCCTTAAAGATCCCTGTGCGGTACCTCGTGAATGCTGACCTGAGCCAGGT 720
DB 661 AACTCCAGTTCCTTAAAGATCCCTGTGCGGTACCTCGTGAATGCTGACCTGAGCCAGGT 720
QY 721 CTAACTTCGTCCTTTTGAAGAAATCAAGTTTCTTACTGTATGATCTTCTTACAGATGG 780
DB 721 CTAACTTCGTCCTTTTGAAGAAATCAAGTTTCTTACTGTATGATCTTCTTACAGATGG 780
QY 781 ATAGCTCAAAAGTTGAAGACACATTTGTAGCAAAACCTGCGCAAAAGAAAGAAACAAG 840
DB 781 ATAGCTCAAAAGTTGAAGACACATTTGTAGCAAAACCTGCGCAAAAGAAAGAAACAAG 840
QY 841 CTTCCTTCAAACTGTTCCTGCTTACCATGTCTTCAACATCAAGTCCATGGGGTAACT 900
DB 841 CTTCCTTCAAACTGTTCCTGCTTACCATGTCTTCAACATCAAGTCCATGGGGTAACT 900
QY 901 TCCAGTCTTACTTCAAGTCTTCCGCAAGACCAATTCCAAATGTGTGTTTCCATAAAGGC 960
DB 901 TCCAGTCTTACTTCAAGTCTTCCGCAAGACCAATTCCAAATGTGTGTTTCCATAAAGGC 960
QY 961 TCCGCAAAATCGCTGGAACCTGCAATTCGAGACCTTAATGCAAGCCTTACACCAAGCCTTAAGA 1020

```

```

DB 961 TCCGCAAAATCGCTGGACCTGCAATTTGAGACCTTAATAATGGAAGCCTCACCAAGCCTTAAGA 1020
QY 1021 GGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACAGGCAATTTCAATAATTATAT 1080
DB 1021 GGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACAGGCAATTTCAATAATTATAT 1080
QY 1081 CTCGGTTATGGGTTCTGTAAAGTAAATCTCGGTGAAGAGCCACACCTCTGTCTTGAAGAAC 1140
DB 1081 CTCGGTTATGGGTTCTGTAAAGTAAATCTCGGTGAAGAGCCACACCTCTGTCTTGAAGAAC 1140
QY 1141 ACTGGCAGCTGGAACATCTGCGCTTGGATCTGTCTTCCATAATATCAAGGCTTCTGAGTGA 1200
DB 1141 ACTGGCAGCTGGAACATCTGCGCTTGGATCTGTCTTCCATAATATCAAGGCTTCTGAGTGA 1200
QY 1201 GCACAAAGTAGCTGCAATAAAG 1224
DB 1201 GCACAAAGTAGCTGCAATAAAG 1224

RESULT 2
AAF60715
ID AAF60715 standard; cDNA; 1652 BP.
XX
AC AAF60715;
AT 03-MAY-2001 (first entry)
DT
EX Murine DNase coding sequence.
DE
EE Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
KW infectious disease; ss.
XX
OS Mus musculus.
XX
EN WO200112793-A1.
XX
PD 22-FEB-2001.
XX
PF 01-MAY-2000; 2000WO-JP002893.
XX
PR 17-AUG-1999; 99JP-00230870.
XX
PA (TANU/) TANUMA S.
XX
PI Tanuma S, Shiokawa D;
XX
PI WPI; 2001-218348/22.
DR P-PSDB; AAB72416.
XX
AC Acidic deoxyribonuclease capable of divalent cation-independent cleavage
of DNA under acidic even neutral pH and not inhibited by G-actin, useful
in remedies for cystic fibrosis and for treatment of infectious diseases.
XX
PS Claim 14; Page 49-51; 61pp; Japanese.
XX
The present sequence is the coding sequence of a murine deoxyribonuclease
(DLAD), which is an endonuclease. DLAD is capable of divalent cation-
independent cleavage of DNA under acidic conditions. DLAD can be used as
a substitute for DNase I in treating cystic fibrosis, and is useful in
the prevention and treatment of infectious diseases
XX
SQ Sequence 1652 BP; 494 A; 410 C; 339 G; 409 T; 0 U; 0 Other;

Query Match 99.7%; Score 1219.8; DB 4; Length 1652;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCAGTCCCTGCAATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 61
DB 175 CACAGTCCCTGCAATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 234
QY 62 CAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 121

```


Db 397 AGAAGATGGAACACCAAGGTTTACGTGTGGAACAGAGTTCAAGGTTCTGGCTG 456
Qy 448 ATACACTCTGTTCCCAAGTTTCCCGAGTTC-----ATGGCTATGAGTACCAACCTCG 501
Db 457 ATTCATTCCTCCTCAGTTCTCTCAATTCGGAAGAGGCTATGATTATCCACCCACA 516
Qy 502 GGAGGCGATATGGAACAAACCGGCTCTGATCAGTTCGGATACAGCCAGTTTGAGGA 561
Db 517 GGAGAGCGAAATGGACAAAGTGGCATCTGCATAACTTCAAGTACCAACAGATGAGGCA 576
Qy 562 ATAGATTTTCAGCTCTTGTGTTTACACCAACCAACATCTACAGCTGCTTCATTCCAAGCACC 621
Db 577 ATAGATTTCTCAGCTCTTGTGTTTACACCAACCAACATCTATAGCTGCTCCATCCAGCCACC 636
Qy 622 TTTCACTGGAACCTTATCTACATGCCCCCGGATGTGCGCAACTCCAGTTCCTTAAAGATC 681
Db 637 TTTCAACGAGGCTCATTCACATGCCCGCCAGCTGTGCACAGGCGCCAGCTCATCAGAGATT 696
Qy 682 CTTGTCGGGTACCTCGTGAACCTGCACTCAGCCAGGCTCTAAGCTTCGTCCTCAATTTGCA 741
Db 697 CTTGGCAGGCTCTCCACACACTTCAGTCGGCCGAGGACAAAAATTCCTCCATTTGCA 756
Qy 742 AAATCAAGTTTATATATGATGATCTTTACAGATGGATAGTCTCAAAAATTTGAAGACA 801
Db 757 AAGTCGGATCTTTCTTGAGCACTCTTTGCGCCCTGGATGCTCAACGGCTGAAGACA 816
Qy 802 CATTTGTTAGCACAAACCTCGCAGAAAAGAAACAGAGCTTCCTCAAACTGTTCCCTG 861
Db 817 CACTTTGTTAACAGAAACCTCGCAGCGAAAAGAAACAGAGCTTCCTCAAACTGTTCCCTT 876
Qy 862 CTTTACATGTTTACACATCAAGTCCATTTGGGTAACTTCCAGTCTTACTTCACTTCT 921
Db 877 CTTTACATGTTTACACATCAAGTCCATTTGGGTAACTTCCAGTCTTACTTCACTTCT 936
Qy 922 CGCAACAGCATTCCTCAATGTTGTTTCCATAAGGCTTCGCAAAATCGTGACCTGC 981
Db 937 TATCAAGATCAGCGCAAGTGTGTATTTCCAAAGGGCCACAAATCGTGACATGT 996
Qy 982 ATTGGAGACCTTAATCGAAGCTTACACCAAGCTTAAAGAGTGGAGGATTCATCTGACA 1041
Db 997 ATTGGAGACCTTAATCGAAGTTCACACCAAGCTTCAAGAGTGGAGGATTCATTTGACC 1056
Qy 1042 AAGATCACAATTTACCGGCAATTTTCAATAATATATCTCCGTTGTTGGTCTGTAAG 1101
Db 1057 CAGAAATTGGCAAAATTACCAAGCAATTTCAAGGATTTAGTATATATGAAAGCTGTAAG 1116
Qy 1102 TAAACTCGGTGAAGGCCACACCTCTGTCCTTGAAACACTGGCAGCTGGAAACATCTGCG 1161
Db 1117 TAAACTGGTGAAGGACACAGGTAATCTATCATTTGAAACCTTGACAAATGGGTCTTCTCC 1176
Qy 1162 CTTGGATCTGTTTCCCAATA 1183
Db 1177 ATTACACCTTCTTATATTTTA 1198

RESULT 6
ID ABK92136
XX ABK92136 standard; DNA; 1140 BP.

XX AC ABK92136;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated DNA sequence #22.

XX KW prostate cancer; prostate tumour tissue; human; mammal; cytostatic;

XX KW gene therapy; gene; ds.

XX OS Mammalia.

XX PN WO200230268-A2.

PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US032045.
XX PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX (BOSB-) EOS BIOTECHNOLOGY INC.
PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX PI P-PSDB; ABG61821.
XX PR WPI; 2002-471335/50.

DR DR
XX PR Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.

XX Claim 22; Page 317; 436pp; English.

PS The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences

XX SQ Sequence 1140 BP; 357 A; 258 C; 245 G; 280 T; 0 U; 0 Other;

Query Match 53.4%; Score 653.4; DB 6; Length 1140;
Best Local Similarity 75.9%; Pred. No. 2e-199;
Matches 836; Conservative 0; Mismatches 256; Indels 9; Gaps 2;

Qy 31 AGATAGAAATGACAGCAAGCCTTAAGACAGTCTTTCTTGTCTTCTTTGCTCTTCTTGGCCCTC 90
Db 37 AATCAGAAATGATGGCAAGACTGCTAAGAACATCTTTGCTTTCTTCTTCTTCTTCTTCTTCTTCT 96
Qy 91 TCTGGGTCTCTGGGACACCAAGAAATCTCATGCAAAATGAATATGTTGAAGCTGTGGAC 150
Db 97 TTTGGGTCTCTGGGACACCAAGAAATCTCATGCAAAATGAAGAAAGGAAAGCTGTGGAC 156
Qy 151 TGGTTATCTTTTATAGTTACCCAAAGGACTAGCAAGGCAAGTGAAGAGCGCGGCTG 210
Db 157 TGGTTATCTTTTATAGTTACCCAAAGGAAAGCTTTTGGTCTTTGCTTTCTTCTTCTTCTTCTTCT 216
Qy 211 CAGTACCTGTACCTGGACTCCACAGACAAACCTTGGAAACAGAGCCTTACTGATTAAC 270
Db 217 GAGTACCTGTACCTGGACTCCACAGAGCTTGGAGGAGAGTGAAGAACTAATGAAT 276
Qy 271 AGCACCAGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATTAATCCAGC 330
Db 277 GACACCAAGAGTGTCTTGGGAAGGACATTAACACAGCTATATGAAGCATATGCTCTTAAG 336
Qy 331 ---AATGACACAGCCTATCTAATATACACGATGTTGCTCCCTGGATCTGTGTAATTACAGC 387

Db 337 AGTAAACACAGCCTATCTTAATATACAAATGATGAGTCCCTAAACCTGTGAATACAGT 396
 Qy 388 AGACAGTATGACATGCAACAGGTTCTGTGTATGGAACAGACGACGAGGTTCTGGCTG 447
 Db 397 AGAAGATGACACACACAAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456
 Qy 448 ATACACTCTGTTCCTCAAGTTTCCCTCCAGTTTC-----ATGGCTATGATGATCCCAACCTCG 501
 Db 457 ATTCATTCCATCCCTCAGTTTCTCCATTCCTCCAAATCCGGAAGAGGCTATGATTAATCCACCCACA 516
 Qy 502 GGGAGGGATATGACACAAACCGGATCTGCAATCACTTTTGGATACACCCAGTTTGGAGAA 561
 Db 517 GGGAGACGAATGGACAAAGTGGCATCTGCATCACTTTTCAAGTACAAACAGTATGAGGCA 576
 Qy 562 ATAGATTTTTCAGCTCTTGGTCTTACACCAACCAACATCTACAGCTGCTTCAATCCAGACAC 621
 Db 577 ATAGATTTTCAGCTCTTGGTCTGCAACCCCAACCTCTATAGCTGCTCCATCCAGCCACC 636
 Qy 622 TTTTCACTGGAACTTATCTATCATGTCGCCCGGATGTGTGCCAACTCCAGTTCTTTAAAGATC 681
 Db 637 TTTTCACTGGAACTTATCTATCATGTCGCCCGGATGTGTGCCAACTCCAGTTCTTTAAAGATC 696
 Qy 682 CCGTGGGATGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 741
 Db 697 CCGTGGGATGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 756
 Qy 742 AAATCAAGTTTATCTATCTATCATGTCGCCCGGATGTGTGCCAACTCCAGTTCTTTAAAGATC 801
 Db 757 AAGTGGATTTCTTTCTTGAAGCACTTTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCA 816
 Qy 802 CAATTTGTTAGCAAAACCTGGCGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 861
 Db 817 CAATTTGTTAGCAAAACCTGGCGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 876
 Qy 862 CCTTACCATGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCT 921
 Db 877 CCTTACCATGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCT 936
 Qy 922 CGCCAGAGCACTTCCAAATGCTGTGTTCCTATTAAGGGCTCCGCAATCGCTGGACCTGC 981
 Db 937 TATCAAGATCAAGCACTGCTGTGTGTATTTCCAAAGGGCACCACAAATCGCTGGACATGT 996
 Qy 982 ATGGAGACCTAAATGGAAGCCCTACACCAAGCTTAAAGAGTGGAGGATTCATCTCTACA 1041
 Db 997 ATGGAGACCTAAATGGAAGCCCTACACCAAGCTTAAAGAGTGGAGGATTCATCTCTACA 1056
 Qy 1042 AAGATCACTACATTTACAGGATTTCAATAATATATCTCCGTTATGGGTTCTGTAAAG 1101
 Db 1057 CAGAATTGGCAAAATTTACCAAGCACTTCAAGGATTTAGTATTATCTATGAAAGCTGTAAG 1116
 Qy 1102 TAAACTCGGTGAAAGGCCACA 1122
 Db 1117 TAAACTCGGTGAAAGGCCACA 1137

RESULT 7
 AAF60716
 ID AAF60716 standard; cDNA; 1086 BP.
 XX
 AC AAF60716;
 XX
 DT 03-MAY-2001 (first entry)
 XX
 DE Human DNase coding sequence.
 XX
 KW Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
 KW infectious disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200112793-A1.
 XX

PD 22-FEB-2001.
 XX
 PF 01-MAY-2000; 2000WO-JP002893.
 XX
 PR 17-AUG-1999; 99JP-00230870.
 XX
 PA (TANU/) TANUMA S.
 XX
 PI Tanuma S, Shiohawa D;
 XX
 DR WPI; 2001-218348/22.
 DR P-PSDB; AAB72417.
 XX
 PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.
 XX
 PS Claim 16; Page 53-55; 61pp; Japanese.
 XX
 CC The present sequence is the coding sequence of a human deoxyribonuclease
 CC (DLAD), which is an endonuclease. DLAD is capable of divalent cation-
 CC independent cleavage of DNA under acidic conditions. DLAD can be used as
 CC a substitute for DNase I in treating cystic fibrosis, and is useful in
 CC the prevention and treatment of infectious diseases
 XX
 SQ Sequence 1086 BP; 339 A; 250 C; 228 G; 269 T; 0 U; 0 Other;
 Query Match 52.4%; Score 641.8; DB 4; Length 1086;
 Best Local Similarity 75.9%; Pred. No. 1.1e-195;
 Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;
 QY 31 AGATAGAAAATGACACAAAGCTCTAAGAACAGTTCTTTTGTCTCTTTTGTCCCTC 90
 Db 4 AAACAGAAAATGATGGCAAGACTGCTAAGAACATCTCTTGTCTTCTTCTTGTCCCTC 63
 QY 91 TCTGGGCTCTCTGGGACACCAAGATCTCATGTCAGAAATGAATATGTTGAAGTGTGGAC 150
 Db 64 TTTGGGCTCTGGGACACCAAGATCTCATGTCAGAAATGAAGAAAGGAAAGCTGTGGAC 123
 QY 151 TGGTTTATCTTTTATAAGTTACCCAAAGGACTAGCAAGGCAAGTGAAGAGGGGGCTG 210
 Db 124 TGGTTTATCTTTTATAAGTTACCTTAAAGACAAACAGAAAGTGGAGAGACTGGTTA 183
 QY 211 CAGTACCTGTACTGACCTCCACAGACAAACCTGGAACAGAGCTCTACCTGATTAAC 270
 Db 184 GAGTACCTGTACTGACCTCTACAACTAGAGCTGGAGGAGAGTGAACAACTAATGAAT 243
 QY 271 AGCACAGGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTTCCAG 330
 Db 244 GACACAGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTTCCAG 303
 QY 331 ---AATGACAGAGCTCTAATATACAAACGATGTCTCTGATCTGTGAATACAGC 387
 Db 304 AGTAAACACAGAGCTCTAATATACAAATGATGAGTCCCTTAAACCTGTGAATACAGC 363
 QY 388 AGACAGTATGACATGCCAAAGTGTCTGTGTATGGAACAGACGAGGTTCTGGCTG 447
 Db 364 AGAAAGTATGACACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTTCTGGCTG 423
 QY 448 ATACACTCTGTTCCTCAAGTTTCCCTCCAGTTTC-----ATGGCTATGATGATCCCAACCTCG 501
 Db 424 ATTCAATCCATCCCTCAGTTTCTCCAAATCCGGAAGAGGCTATGATTAATCCACCCACA 483
 QY 502 GGGAGGGATATGACACAAACCGGATCTGCACTCACTTTTCGGATACAGCAGTTTTCAGGAA 561
 Db 484 GGGAGACGAATGGACAAAGTGGCATCTGCATCACTTTCAAGTACCAACAGTATGAGGCA 543
 QY 562 ATAGATTTTTCAGCTCTTGGTCTTCAACCAACATCTACAGCTGCTTCAATTTCCAGACAC 621
 Db 544 ATAGATTTTTCAGCTCTTGGTCTGCAACCCCAACCTCTATAGCTGCTCCATCCAGCCACC 603
 QY 622 TTTCACTGGAACCTTATCTATCATGCTCCCGGATGTGTGCCAACTCCAGTTCTTTAAAGATC 681

Db 604 TTTCACGAGCTCATTACATGCCCCAGCTGTGACACGAGGCCAGCTCATCAGAGATT 663
QY 682 CCTGTCGGTACCTCCTGACCTGACCTGACGCCAGGGTCTAACTTCGTCGCTTTTGCA 741
Db 664 CTTGCGAGGCTCTCACCACCTTCAGTCGGCCGAGGACAAAAATTCCTCCATTTTGCA 723
QY 742 AATCAAGATTTTATCTAGTACGACCTTTACAGGATGGATGCTCAAAAGTTTGAAGACA 801
Db 724 AAGTGGATTTCTTTCTGATGCACTTTTGCAGCTGGATGCTCAACGGCTGAAGACA 783
QY 802 CATTTGTTAGCACAACTTGGCAGAAAAAAGAAACAGAGCTTCCTTCAAACTGTTCCCTG 861
Db 784 CATTTGTTACAGAACTTGGCAGGAAAGAAAGACAGAGCTTCCTTCAAACTGCTCCCTT 843
QY 862 CTTTACATGTCACAACTCAAGTCAATTTGGGGTAACTTCCAGTCTTACTTTCAGTTCT 921
Db 844 CTTTACATGTCACAACTCAAGTCAATTTGGGGTAACTTCCAGTCTTACTTTCAGTTCT 903
QY 922 CGCCAGACCAATTCCTCAAGTGTGTTTCCATAAAGGCTCCGCAATCGCTGACCTGC 981
Db 904 TATCAAGATCATGCCAAGTGTGTTTCCATAAAGGCTCCGCAATCGCTGACATGT 963
QY 982 ATTGGAGACCTTAATCGAAGCTACACCAAGCCTTTAGAGGTGGAGGATTCATCTGTACA 1041
Db 964 ATTGGAGACCTTAATCGAAGTCCACACCAAGCCTTCAGAAAGTGGAGATTCATTTGACC 1023
QY 1042 AGAATCCTACATTTACCGAGGCAATTTCAATAATTAATATCTCGTTATGGTCTGTAAAG 1101
Db 1024 CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTAGTATTATATCTAGAAAGCTGTAAAG 1083
QY 1102 TAA 1104
Db 1084 TAA 1086
RESULT 8
ACH19937
ID ACH19937 standard; cDNA; 480 BP.
XX AC ACH19937;
XX XX
DT 13-OCT-2003 (first entry)
XX DE Human adult lung cDNA #940.
XX XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX XX
FN US2003073623-A1.
PD 17-APR-2003.
XX XX
PF 30-JUL-2001; 2001US-00918995.
XX XX
PR 30-JUL-2001; 2001US-00918995.
XX XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX XX

PS Claim 1; SEQ ID NO 7149; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX SQ Sequence 480 BP; 134 A; 120 C; 106 G; 118 T; 0 U; 2 Other;
Query Match 15.8%; Score 193.2; DB 8; Length 480;
Best Local Similarity 79.2%; Pred. No. 3e-51;
Matches 244; Conservative 0; Mismatches 58; Indels 6; Gaps 1;
QY 331 AATGACACAGCCTATCTAATATATACACGATGGTGTCCCTGGATCTGTGAATTACAGCAGA 390
Db 171 AACACACACGCTATCTAATATATACATGATGGATGCTCCCTAAACCTGTGAATTACAGCAGA 230
QY 391 CAGTATGGACATGCCAAGGTCTGTGGTATGGAAACAGAACGCGAGGGTCTTGCTGATTA 450
Db 231 AAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTGATT 290
QY 451 CACTCTGTTCCCAAGTTTCCCTCCAGTTC-----ATGCGTATAGTACCCAACTCCGGG 504
Db 291 CATTCCATCCCTCAGTTTCTCTCCAAATTCGGAAGAGGCTATGATATCCACCACAGGG 350
QY 505 AGCGATATGACAAAACCGGCATCTGCATCACTTTCGGATACAGCCAGTTTGAGGAATA 564
Db 351 AGAGAAATGGACAAAGTGGCATCTGCATACTTTCAAGTACACACCATGATGAGCAATA 410
QY 565 GATTTTCAGCTCTTGGTCTTTACAAACCAACATCTACAGCTGCTTCATTCCCAAGCACCTTT 624
Db 411 GATTCTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGTCTCCATCCCGAGCACCTTT 470
QY 625 CACTGGA 632
Db 471 CACCAGGA 478
RESULT 9
ADE87477/C
ID ADE87477 standard; DNA; 266145 BP.
XX AC ADE87477;
XX XX
DT 29-JAN-2004 (first entry)
XX XX
XX Fowlpox virus genome DNA.
XX fowlpox virus; FPV; virucide; tuberculostatic; protozoacide; antipretic;
KW cytotatic; hepatotropic; antibacterial; vaccine; malaria; tuberculosis;
KW East Coast fever; avipox virus; influenza; hepatitis;
KW human papilloma virus; tumour; leishmaniasis; listeriosis; theileria;
KW gene; ds.
XX OS Fowlpox virus.
XX XX
PN WO2003047617-A2.
XX XX
PD 12-JUN-2003.
XX XX

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
 CC)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridizing selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini-
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 65 BP; 15 A; 22 C; 10 G; 18 T; 0 U; 0 Other;
 SQ
 Query Match 5.3%; Score 65; DB 6; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1118 CCACACCTTCCTTGAACACTGCGACTGGAACATCTCGCTGGATCTGTTCTCC 1177
 DB 1 CCACACCTTCCTTGAACACTGCGACTGGAACATCTCGCTGGATCTGTTCTCC 60
 QY 1178 ATAT 1182
 DB 61 ATAT 65
 RESULT 11
 AAV02903
 ID AAV02903 standard; DNA; 1575 BP.
 AC AAV02903;
 XX 08-MAY-1998 (first entry)
 DT Human DNase II DNA.
 DE
 XX DNase II; human; deoxyribonuclease; pulmonary disease; diagnostic;
 KW cellular DNA debris; ss.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT misc_feature 29..31 /tag= a
 FT /note= "Region indicated in specification"
 FT CDS 86..1168 /tag= b
 FT /product= "DNase II"
 FT sig_peptide 86..133 /tag= c
 FT 134..1165 /tag= d
 FT mat_peptide /product= "DNase II"
 FT
 XX WO9740134-A2.
 PN
 XX 30-OCT-1997.
 PD
 XX 23-APR-1997; 97WO-US006664.
 PF

XX 25-APR-1996; 96US-00639294.
 PR (GETH) GENENTECH INC.
 XX Baker KP, Baron WF;
 PI WPI; 1997-535820/49.
 XX P-PSDB; AAW31495.
 DR Human DNase II and related nucleic acids - useful in protein production
 XX e.g. for therapeutic use to treat systemic lupus erythematosus and
 PT pulmonary diseases e.g. cystic fibrosis.
 PS Claim 1; Fig 1; 30pp; English.
 XX This sequence encodes a novel human deoxyribonuclease, DNase II. This
 CC protein is useful to reduce the viscoelasticity of DNA-containing
 CC material, e.g. mucus, and used to treat patients with pulmonary diseases
 CC or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.
 CC It may also improve antibiotic efficacy in the treatment of abscesses,
 CC infected lesions etc. provide treatment in non-infected conditions in
 CC which there is an accumulation of cellular DNA debris, e.g.
 CC pyelonephritis and be used to degrade DNA in biological samples or in
 CC diagnostic assays. The nucleic acids and vectors can be used for in vivo
 CC or ex vivo gene therapy, and antisense oligonucleotides can be produced
 CC from the nucleic acids, which can bind to and prevent expression of
 CC nucleic acid within cells. The nucleic acids (or a portion) can also be
 CC used for hybridisation assays for nucleic acids encoding human DNase II
 CC in a sample, or to identify and isolate nucleic acids sharing substantial
 CC sequence identity (e.g. encoding naturally-occurring allelic variants of
 CC human DNase II). The antibodies can be used to detect and measure human
 CC DNase II in tissues or clinical samples, and in the purification of human
 CC DNase II. The nucleic acids enable production of human DNase II by
 CC recombinant DNA methods in quantities sufficient for clinical use, not
 CC previously possible
 XX Sequence 1575 BP; 358 A; 465 C; 450 G; 302 T; 0 U; 0 Other;
 SQ
 Query Match 5.0%; Score 60.8; DB 2; Length 1575;
 Best Local Similarity 44.6%; Pred. No. 2.9e-08;
 Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;
 QY 200 AGCGGGGCTGCAGTACTGTACTGCACTCCACAGACAACTCTGGAACAGAGCCCTCT 259
 DB 222 AGAGAGGGCTGCAGTACTGCACTCCAGTATCTGACGAGAGCTCGGAGGCTGGCGGACGCGAGGG 281
 QY 260 ACCTGATTAACAGCACGAGGAGTCTCTGGGAGGAGCTTACAGCATCTGTATGACACAC 319
 DB 282 CACTCATCAACAGCCCGAGGGGGCGTGGGCGGAGGCTGACGCGCTGTACCGGAGCA 341
 QY 320 ATATATCCAGATGACAGACGCTTATCTATATACAGATGGTGTCTCCTGGATCTGTGA 379
 DB 342 ACACAGGAGCTGCTCTCTCTCTACATGACCAACCGCTTCAACCCAGAGGCTC 401
 QY 380 ATTACAGCAGACAGTATGGACATGCCAAAGTCTGTGGTATGGAACAGAGCGAGGGT 439
 DB 402 AGGACTCTTCCATGCTGGGACACAGAGGGTGTCTGCTCTCTTACACGATGGGGCT 461
 QY 440 TCTGGTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATGGTATGAGTACCAACT 499
 DB 462 TCTGGTGGTCCACAGTGTACTTAACCTTCCCTCCACCGGCTCTCTCTCTGATACAGCT 521
 QY 500 CG-----GGGAGGCGATATGGACAAACCGGATCTGATCATCACTTTCGGATACAGCC 550
 DB 522 GGCCTCATAGCGCTGTACCTACGGGACAGCCCTCTGTGTGTCTTTTCCCTTCGGCTC 581
 QY 551 AGTTTGAGGAATAGATTTTTCAGCTCTTGGTCTTACACCAACATCTACAGCTGCTTCA 610
 DB 582 AGTTCTCGAAGATGGGCAAGCAGCTGACCTACACCTACCCCTGGGTCTATAATACAGC 641
 QY 611 TTCCAGCACCTTTCACCTGGAACTTATCTACATGCCCGGATGTGTGCCACTCCAGTT 670

Db 642 TGGAGGGATCTTTGGCCAGAAATTCGCCGACTTGGAGAAATGTGGTCAAGGGCCACCACG 701
 QY 671 CCTTAAAGATCCCTGTCCGCTACCTGCACTGCACTGAGCCAGGCTTAAACTTCG 730
 Db 702 TTAGCCAGAACCTTGAACAGAGCATCACATCCAGGCCGGGCTGTTTTC 761
 QY 731 TCCATTTGCAAAATCAAGTTTTTATCTGATGACATCTTTACAGGATGATAGCTCAAA 790
 Db 762 AGAGCTTTGCCAAGTTCCAGAAATTTGGAGATGACCTGTACTCCGGCTGTTGGCAGCAG 821
 QY 791 AGTTGAAGACATTTGTTAGCAACAACCTGGCAGAAAAGAAACAAGAGCTTCTTCAA 850
 Db 822 CCTTGTGTAACACCTGCAGTCCAGTTCCTGGCACAACAACTGTAGGCATCTCGCCTCTA 881
 QY 851 ACTGTTCCCTGCCTTACCATGTTCAACAATCAAGTCCATTGGGT-----ACATTCCA 904
 Db 882 ACTGCTCGGATATCTGGCAGTTCTGAATGTGAACAGATAGCTTCCCTGGACCCAGCCG 941
 QY 905 AGTCTTACTTCACTTCTGCGCAAGACATTTCCAAATGGTGTGTTTCCATAAAGGGCTCCG 964
 Db 942 GCCCAAGCTTCAACAGCAGAGGACCACTCCAAATGGTGTGTTGCCCAAAAGG----- 996
 QY 965 CAAATCGCTGGACCTGCATTGGAGACCTAAATCGAAGCCTACACCAAGCTTAAAGAGTG 1024
 Db 997 ----GCCCTGGACCTGCGTGGGTGACATGAATCGGAACCAAGGGAGAGGACCAACGGGGTG 1052
 QY 1025 GAGGATTCATCTGTCAAAAGAAATCACTATACATTACCAGGCATTTC 1070
 Db 1053 GGGGCACACTGTGTGCCAGCTGCCAGCCCTCTGGAAGCCCTTCCA 1098

RESULT 12

AAV29137
 ID AAV29137 standard; cDNA; 1915 BP.
 XX
 AC AAV29137;
 XX
 DT 11-SEP-1998 (first entry)
 XX
 DE Homo sapiens DNase II gene.
 XX
 KW DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
 KW controlled cell death; apoptosis; metamorphosis; cell turnover;
 KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
 KW autoimmune disease; diagnosis; anticancer agents;
 KW enhanced chromosomal rearrangement; chromosome instability; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 68..1150
 FT /*tag= a
 FT /product= "deoxyribonuclease (DNase II)"
 XX
 PN W09816659-A1.
 XX
 PD 23-APR-1998.
 XX
 PF 09-OCT-1997; 97WO-US018262.
 XX
 PR 15-OCT-1996; 96US-0028539P.
 XX
 PA (DART-) DARTMOUTH COLLEGE.
 XX
 PI Eastman A, Krieser R;
 XX
 XX WPI; 1998-251301/22.
 DR P-PSDB; AAW37920.
 XX
 PT Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
 PT e.g. to induce apoptosis in tumour cells, identify agents modulating
 PT apoptosis and digest DNA in sputum of cystic fibrosis patients.
 XX

PS Claim 2; Page 15-16; 29pp; English.

XX The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme
 CC is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis
 CC patients to reduce sputum viscosity. It (or fragments) are also useful to
 CC produce antibodies, e.g. to study DNase II expression in cells. DNase II
 CC has recently been linked with DNA fragmentation in the early stages of
 CC controlled cell death (apoptosis), a process critical to homeostasis
 CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead
 CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
 CC whilst too little can lead to cancer or autoimmune diseases. The
 CC antibodies raised against it can therefore be used to diagnose apoptotic
 CC stages in selected cells, by contacting cells with the antibody,
 CC detecting binding of the antibody with DNase II and determining DNase II
 CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic
 CC agents, e.g. anticancer agents to promote apoptosis in cells. The
 CC antibodies can also be used to identify cells susceptible to premature
 CC death. The cDNA is useful to identify agents modulating apoptosis in
 CC cells, by treating cells with an agent, transfecting cells with cDNA and
 CC monitoring apoptosis compared with untreated cells. Inhibitors identified
 CC may be useful in preventing diseases relating to enhanced chromosomal re-
 CC arrangement. Vectors comprising the cDNA can be used to induce apoptosis
 CC in selected cells, e.g. tumour cells or cells involved in autoimmune
 CC disorders. Antisense oligonucleotides can be administered to cells to
 CC inhibit DNase II expression to reduce chromosome instability associated
 CC with cancer
 XX

Sequence 1915 BP; 462 A; 533 C; 530 G; 390 T; 0 U; 0 Other;

Query Match 5.0%; Score 60.8; DB 2; Length 1915;
 Best Local Similarity 44.6%; Pred. No. 3.3e-08;
 Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;
 QY 200 AGCGGGGCTGCAGTACTGTGTACCTGCACTCCACAGACAAACCTGGAACAGAGCTCT 259
 Db 204 AGAGAGGGCTGCAGTACAAAGTATCTGGACAGAGCTCCGGAGGCTGGCGGCGGAGGG 263
 QY 260 ACCTGATTAACAGACACAGGAGTGTCTTGGGAGGACCTTACAGATCTGTATGACACAC 319
 Db 264 CACTCATCAACAGCCCGGAGGGGGCGGTGGGCCGAAGCTGCGAGCGCTGTACCGGAGCA 323
 QY 320 ATAATTCCACGAATGACACAGCCTATCTAATATACAGATGGTGTCCCTGGATCTGTGA 379
 Db 324 ACACAGCAGCTCGCTTCTGTCTCTACATGACCAACCGCTCAACCCAGCAAGGCTC 383
 QY 380 ATTACAGCAGACAGTATGACATGCGCAAGGCTGTCTGTATGGAACAGAACGCGAGGCT 439
 Db 384 AGGACTCTTCCATGCTGGGCACACGAAGGGTGTCTGTCTTGCACCATGCGGGCT 443
 QY 440 TCTGGCTGATACACTCTGTTCCTCAAGTTTCCCGCAGT-----TCATGGCTATGAGT 490
 Db 444 TCTGGCTGGTCCACAGTGTACCTAACTTCCCTCCACCGGCTCTCTGTGATACAGTT 503
 QY 491 ACCCAACCTCGGGAGGCGGATATGACAAAACCGGCATCTGCATCACTTTCGGATACAGCC 550
 Db 504 GGCCTCATAGCGCTGTACCTACGGGCGAGACCTGCTCTGTGTGTCTTTCCTTCTGCTC 563
 QY 551 AGTTTGAGGAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACAGCTGCTTCA 610
 Db 564 AGTTCTCGAAGATGGGCAAGCAGCTGACCTACACCTACCCCTGGGCTCTATACTACCAGC 623
 QY 611 TTCCNAGCACCCTTTCACCTGGAACCTTATCTACATGCCCCGGATGTGTGCCAATCCTCAGTT 670
 Db 624 TGGAAGGGATCTTTGCCCGAGGAATTCCTCCAGTGTGGAGAATGTGGTCAAGGCCACACG 683
 QY 671 CCTTAAAGATCCCTGCTCCGGTACCTCGCTGACCTGACCTCAGCCAGGCTTAACATTGG 730
 Db 684 TTAGCCAGAACCTTGGAAACAGCAGCATCACTACATCCACAGCCCGGGCTGTTTTCC 743
 QY 731 TCCATTTTGGCAAAATCAAGTTTTTATCTGATGACATCTTTTACAGGATGGATAGTCAAA 790
 Db 744 AGAGCTTTGCCAAGTTCAGCAAAATTTGGAGATGACCTTACTCCGGCTGTGTGGACAG 803

QY 791 AGTTGAGACACATTTGTTAGCACAACTGGCAGAAAAGAAACAGAGCTTCCTTCAA 850
 DB 804 CCCTTGGTACCAACCTGCGAGTCCAGTTCTGGCACAATACTGAGCATCTCCTCTTA 863
 QY 851 ACTGTTCCCTGCTTACCATGCTTACAACATCAAGTCCATFTGGGGT-----AACTTCCA 904
 DB 864 ACTGCTCGATATCTGGCAGTTCTGATGTGACCAAGATAGCTTCCCTGGACAGCGG 923
 QY 905 AGTCTTACTTCAGTTCTGCGCAAGACCATTCGAAATGGTGTGTTTCCATAAAGGGCTCCG 964
 DB 924 GCCCAAGCTTCAACAGCACAGAGGACCATCTCCAAATGGTGTGTTCCCAAAAGG----- 978
 QY 965 CAAATCGCTGACCTGATTTGGAGACCTAAATCGAGCCCTACACCAAGCCTTAAGAGGTG 1024
 DB 979 -----GCCCTGGACCTGCTGGTGGGTGACATGAATCGAACCGAGGAGGACCAAGGGTG 1034
 QY 1025 GAGGATTCATCTGTACAAAAGATCACTACATTTACCAGGCAATTTCA 1070
 DB 1035 GGGGCACACTGTGTGCCAGCTGCCAGCCCTCTGGAAAGCCTTCCA 1080

RESULT 13
 ABL13359
 ID ABL13359 standard; cDNA; 1201 BP.

XX AC ABL13359;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 34559.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ss.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX P-PSDB; ABB69256.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PT interactions.

XX PS Claim 1; SEQ ID NO 34559; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
 XX CC useful in developmental biology and in elucidating cell signalling and
 XX CC cell-cell interactions in higher eukaryotes for the development of
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
 XX CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 XX CC ABBS72072). The sequence data for this patent did not form part of the
 XX CC printed specification, but was obtained in electronic format directly
 XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1201 BP; 300 A; 302 C; 321 G; 278 T; 0 U; 0 Other;
 Query Match 4.9%; Score 59.4; DB 4; Length 1201;
 Best Local Similarity 55.6%; Pred. No. 7.1e-08;

Matches 114; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
 QY 396 TGGACATGCCAAGGTCCTGCTATGGAACAGACGACAGGGGTTCTGCTGATACACTC 455
 DB 366 TGGTCAATGCCAAGGAGTGTGTCCTGATGCGGAGACGGCTATTGATTTGATTTCCACTC 425
 QY 456 TGTTCCTCAAGTTTCCCTCCAGTTTCATGGCTATGAGTACCCCAACCTCGGGAGGCGATANGG 515
 DB 426 GTGTGCCCAATTTCCCAACAATTCGGGACTACAGCTATCCACATCCGGCGAGCAGTACGC 485
 QY 516 ACAAACCGGCATCTGATCATCTTTCGATACACCGAGTTTGAGGAAATAGATTTTCAGCT 575
 DB 486 ACAGAGCATGCTGTGGTGCACACTGAAGGCGAGGATTTGGAGAGGTTGTCAGATTCT 545
 QY 576 CTGTGCTTTACACCAACCAATCTAC 600
 DB 546 GGTGTACATGAACCGCATTCTAC 570

RESULT 14
 ABL13358
 ID ABL13358 standard; cDNA; 3760 BP.

XX AC ABL13358;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 34556.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ss.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX P-PSDB; ABB69255.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PT interactions.

XX PS Claim 1; SEQ ID NO 34556; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
 XX CC useful in developmental biology and in elucidating cell signalling and
 XX CC cell-cell interactions in higher eukaryotes for the development of
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
 XX CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 XX CC ABBS72072). The sequence data for this patent did not form part of the
 XX CC printed specification, but was obtained in electronic format directly
 XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 3760 BP; 1214 A; 732 C; 739 G; 1075 T; 0 U; 0 Other;
 Query Match 4.9%; Score 59.4; DB 4; Length 3760;
 Best Local Similarity 55.6%; Pred. No. 1.4e-07;
 Matches 114; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 396 TGGACATGCGCAAGGTCGTCTGATGGAACAGAACAGCGGGTCTCGGTGATACACTC 455
 Db |||||
 QY 1925 TGGTCATGCGCAAGGAGTGGTGGCCAGTATGGGAGACGSCATTTGGATTGTCACACTC 1984
 Db |||||
 QY 456 TGTTCACAACTTCCCGAGTTCATGCTATGCTATGAGTACCCAACTTCGCGGAGGCGATATGG 515
 Db |||||
 QY 1985 GGTGCCCAAAATTTCCAAATTCGGACTACAGCTATCCACATCCCGGAGCAGTACGC 2044
 QY 516 ACAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGGGAATAGATTTTCAGCT 575
 Db |||||
 QY 2045 ACAGAGATCTCTGCTGCTACACTGAAGGCGAGGATTTGGAGAAGTTGGTCTGATTTCT 2104
 QY 576 CTTGGTCTTACAAACCAACACTCTAC 600
 Db |||||
 QY 2105 GGTGTACATGAACCGCACTTCTAC 2129

RESULT 15

AAV29138
 ID AAV29138 standard; cDNA; 927 BP.

XX AC
 XX AAV29138;

DT 11-SEP-1998 (first entry)

XX Bos taurus DNase II gene.

XX DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;
 KW controlled cell death; apoptosis; metaplasia; cell turnover;
 KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;
 KW autoimmune disease; diagnosis; anticancer agents;
 KW enhanced chromosomal rearrangement; chromosome instability; ss.

XX Bos taurus.

EH Key Location/Qualifiers
 FT 1.831
 CDS /*tag= a

FT /product= "deoxyribonuclease II (DNase II)"

XX WO9816659-A1.

XX 23-APR-1998.

XX 09-OCT-1997; 97WO-US018262.

XX 15-OCT-1996; 96US-0028539P.

XX (DART-) DARTMOUTH COLLEGE.

XX Eastman A, Krieser R;

XX WPI; 1998-251301/22.

XX P-PSDB; AAW37921.

PT Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,
 PT e.g. to induce apoptosis in tumour cells, identify agents modulating
 PT apoptosis and digest DNA in sputum of cystic fibrosis patients.

PS Claim 2; Page 17; 29pp; English.

XX The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme
 CC is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis
 CC patients to reduce sputum viscosity. It (or fragments) are also useful to
 CC produce antibodies, e.g. to study DNase II expression in cells. DNase II
 CC has recently been linked with DNA fragmentation in the early stages of
 CC controlled cell death (apoptosis), a process critical to homeostasis
 CC during, e.g. metaplasia or cell turnover. Too much cell death can lead
 CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),
 CC whilst too little can lead to cancer or autoimmune diseases. The
 CC antibodies raised against it can therefore be used to diagnose apoptotic
 CC stages in selected cells, by contacting cells with the antibody,
 CC detecting binding of the antibody with DNase II and determining DNase II

CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic
 CC agents, e.g. anticancer agents to promote apoptosis in cells. The
 CC antibodies can also be used to identify cells susceptible to premature
 CC death. The cDNA is useful to identify agents modulating apoptosis in
 CC cells, by treating cells with an agent, transfecting cells with cDNA and
 CC monitoring apoptosis compared with untreated cells. Inhibitors identified
 CC may be useful in preventing diseases relating to enhanced chromosomal re-
 CC arrangement. Vectors comprising the cDNA can be used to induce apoptosis
 CC in selected cells, e.g. tumour cells or cells involved in autoimmune
 CC disorders. Antisense oligonucleotides can be administered to cells to
 CC inhibit DNase II expression to reduce chromosome instability associated
 CC with cancer

XX Sequence 927 BP; 214 A; 258 C; 237 G; 218 T; 0 U; 0 Other;

Query Match 4.5%; Score 55.4; DB 2; Length 927;

Best Local Similarity 46.3%; Pred. No. 1.2e-06;

Matches 311; Conservative 0; Mismatches 336; Indels 24; Gaps 3;

QY 396 TGGACATGCGCAAGGTCGTCTGATGGAACAGAACAGCGGGTCTCGGTGATACACTC 455

Db |||||

QY 456 TGTTCACAACTTCCCGAGTTCATGCTATGCTATGAGTACCCAACTTCGCGGAGGCGATATGG 515

Db |||||

QY 135 CGTTCACAACTTCCCGAGTTCATGCTATGCTATGAGTACCCAACTTCGCGGAGGCGATATGG 194

QY 507 GCGATATGGAACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGGGAATAGATA 566

Db |||||

QY 195 AAAATATGGGAGACCTGATCTGTATCTTTCTCTCACCAGTTCTGGATATCAG 254

QY 567 TTTTCAGCTCTTGGTCTTACAAACAAATCTACAGCTGCTTCAATTCAGACACCTTTTCA 626

Db |||||

QY 627 CTGGAACCTTATCTCATATGCTCCCGGATGTGTGCAACTCCAGTTCCTTAAAGATCCCTGT 686

Db |||||

QY 315 CCAGAAATTCCTCTACCTCGAGGAGGTAGTCAAGGCCATCAGTTCCGCGAGGACCGTG 374

QY 687 CCGGTACTCTCGTGAACACTGCATCTCAGCCCGGCTCTAAACTTCGTCTCAATTTTGCAAAATC 746

Db |||||

QY 375 GAACAGCAGTGTAACTCATCATCAATCAAGAAAGAGGCCACATTCAGAGCTTTGCCAATTT 434

QY 747 AAGTTTATATCTGATGATCATCTTTACAGGATGATAGTCTCAAAAGTTGAAGACACATTT 806

Db |||||

QY 435 TGGAAACTTTGGAGATGACCTGTACTCTGCTGGTGGCGGAAGCCCTTGGCAGTACCCCT 494

QY 807 GTTAGCACAAACCTGCGAGAAAGAAACAGAGCTTCTTCAAACTGTTCCTGCTTCA 866

Db |||||

QY 495 GCAGGTCCAAATTCGCAACAGATCTTCTGGATTCCTTCAAACTGTTCCTGCTTCA 554

QY 867 CCATCTCTACACATCAAGTCCATTTGGGGTAACT-----TCCAAAGTCTTACTTCAAGTTC 920

Db |||||

QY 555 GCATGTATTTGACGTGACTCAGACAGCTTTCCTCGGCGCAGCTGGCGCAGCTTCAATGC 614

QY 921 TCGCCCAAGACCAATTCCAATGTGTGTTTCCATTAAGGCTCCGCAAAATTCGTGACCTTG 980

Db |||||

QY 615 CACAGAAAGCCATTCCAAGTGTGTGTAAACCCCAAAAGG-----GCCCTGGGCGCTG 665

QY 981 CATTGGAGACCTTAAATCGAAGCCTTACCAAGCCTTTAAGAGGTGAGGATTCATCTGTAC 1040

Db |||||

QY 666 TGTGGGTGACATGAATCGGAACCAAGAGAGAGCACCAGGGGTGGGGGCACTCTGTGTGC 725

QY 1041 AAGAATCACT 1051

Db |||||

QY 726 CCAGATGCTCT 736

Search completed: October 14, 2004, 01:02:01
 Job time : 470.631 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:33:41 ; Search time 4727.53 Seconds
(without alignments)
11221.906 Million cell updates/sec

Title: US-10-790-589-1

Perfect score: 1224

Sequence: 1 tccagtcctccgtcatgaa.....aacgtagcgtcccaataaag 1224

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1224	100.0	1224	6	AR201143 Sequence
2	1219.8	99.7	1652	6	AR432046 Sequence
3	1219.8	99.7	1652	6	BD091714 Novel deo
4	1219.8	99.7	1654	10	AF128888 Mus muscu
5	867.6	70.9	1498	10	AF178974 Rattus no
6	672.8	55.0	1268	6	AR201144 Sequence
7	653.4	53.4	1140	9	AF274571 Homo sapi
8	641.8	52.4	1086	6	AR432047 Sequence
9	641.8	52.4	1086	6	BD091715 Novel deo
10	512.2	41.8	1029	9	AF333389 Homo sapi
11	460.4	37.6	2337	10	AF334608 Mus muscu
12	460.4	37.6	197457	2	AC127571 Mus muscu
13	460.4	37.6	211003	2	AC141632 Mus muscu
14	460.4	37.6	255925	2	AC114618 Mus muscu
15	308.2	25.2	231698	2	AC118117 Rattus no
16	308.2	25.2	306870	2	AC098557 Rattus no
17	255.6	20.9	19164	9	AF334602 Homo sapi
18	255.6	20.9	164529	2	AC027606 Homo sapi
19	255.6	20.9	205034	9	AL359273 Human DNA
20	202.6	16.6	642	10	AF334603 Mus muscu
21	179	14.6	1374	10	AF334607 Mus muscu
22	160.4	13.1	719	10	AF334603S2 Mus muscu
23	151.4	12.4	1825	10	AF334603S1 Mus muscu
24	114.8	9.4	1403	10	RATSP6T7
25	97.4	8.0	86816	2	ALI36113
26	91.6	7.5	7470	14	FPO6408
27	91.6	7.5	288539	14	AF198100 Fowlpox v
28	84	6.9	1391	9	AB074101 Aotus tri
29	82.4	6.7	1357	9	AB074093 Homo sapi
30	82.4	6.7	1388	9	AB074098 Papio ham
31	82.4	6.7	1390	9	AB074097 Hylobates
32	82.4	6.7	1391	9	AB074099 Macaca mu
33	82	6.7	1089	10	AF334603S3 Mus muscu
34	80.8	6.6	1387	9	AB074094 Pan trogl
35	80.8	6.6	1388	9	AB074096 Pongo pyg
36	80.8	6.6	1390	9	AB074100 Macaca fa
37	79.2	6.5	1388	9	AB074095 Gorilla g
38	63	5.1	42331	3	CBRG45A03 Caenorhab
39	60.8	5.0	1083	9	BT007047 Homo sapi
40	60.8	5.0	1083	12	BT007514 Synthetic
41	60.8	5.0	1575	6	AR337903 Sequence
42	60.8	5.0	1593	9	AB004574 Homo sapi
43	60.8	5.0	1601	9	BC010419 Homo sapi
44	60.8	5.0	1751	9	AF045937 Homo sapi
45	60.8	5.0	1920	9	AF047016 Homo sapi

ALIGNMENTS

RESULT 1	AR201143	Sequence 1	1224 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201143	Sequence 1	from patent US 6358723.			
DEFINITION	AR201143	Sequence 1	from patent US 6358723.			
ACCESSION	AR201143	Sequence 1	from patent US 6358723.			
VERSION	AR201143.1	GI:20252031				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1224)					
AUTHORS	Eastman, A. Richard. and Krieser, R. Joe.					
TITLE	Deoxyribonuclease II beta. proteins and cDNAs					
JOURNAL	Patent: US 6358723-A 1 19-MAR-2002;					
FEATURES	Location/Qualifiers					

Pred. No. is the number of results predicted by chance to have a


```
QY 422 GGAACAGACGAGGGGTTCTGGCTGATACACTCTGTTCCTCAAGTTTCCCCAGTTTCATG 481
Db 595 GGAACAGACGAGGGGTTCTGGCTGATACACTCTGTTCCTCAAGTTTCCCCAGTTTCATG 654
QY 482 GCTATGAGTACCAACCTCGGGGAGGATATGGACAAACCGCATCTGCATCATTTCG 541
Db 655 GCTATGAGTACCAACCTCGGGGAGGATATGGACAAACCGCATCTGCATCATTTCG 714
QY 542 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 601
Db 715 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 774
QY 602 GCTGCTTCATTCACAGCATTTCATCTGAAACCTTATCTACATGCCCCGAGATGTCGA 661
Db 775 GCTGCTTCATTCACAGCATTTCATCTGAAACCTTATCTACATGCCCCGAGATGTCGA 834
QY 662 ACTCCAGTTTCTTAAAGATCCCTGTGCTGGTACCTGCTGAACTGCACTGAGCCAGGTC 721
Db 835 ACTCCAGTTTCTTAAAGATCCCTGTGCTGGTACCTGCTGAACTGCACTGAGCCAGGTC 894
QY 722 TAAACTTCGTTCCATTTTGCAAAATCAAGTTTATATCTGATGACATCTTTACAGGATGA 781
Db 895 TAAACTTCGTTCCATTTTGCAAAATCAAGTTTATATCTGATGACATCTTTACAGGATGA 954
QY 782 TAGCTCAAAAGTTGAAGACACATTTGTTAGCAAAACCTGGCAGAAAAAGAAACAGAGC 841
Db 955 TAGCTCAAAAGTTGAAGACACATTTGTTAGCAAAACCTGGCAGAAAAAGAAACAGAGC 1014
QY 842 TTCCCTTCAAACTGTTCCCTGCTTACCATGCTTACAACTCAAGTCCATTTGGGTTAACTT 901
Db 1015 TTCCCTTCAAACTGTTCCCTGCTTACCATGCTTACAACTCAAGTCCATTTGGGTTAACTT 1074
QY 902 CCAAGTTTACTTCAAGTTTCTGCGCAAGACCAATTCAAATGTTGTTTCCATTAAGGGCT 961
Db 1075 CCAAGTTTACTTCAAGTTTCTGCGCAAGACCAATTCAAATGTTGTTTCCATTAAGGGCT 1134
QY 962 CGGCAATCGTGGACCTGCTGATGAGACCTAAATCGAAGCCTACACCAAGCCTTAAGAG 1021
Db 1135 CGGCAATCGTGGACCTGCTGATGAGACCTAAATCGAAGCCTACACCAAGCCTTAAGAG 1194
QY 1022 GTGGAGGATTCATCTGACAAAGATCACTACATTTACAGGATTTCTAATATATATC 1081
Db 1195 GTGGAGGATTCATCTGACAAAGATCACTACATTTACAGGATTTCTAATATATATC 1254
QY 1082 TCGTTATGGGTTCTGTAAGTAACTCGGTGAAAGCCACACCTCTGTCTTGAATAACA 1141
Db 1255 TCGTTATGGGTTCTGTAAGTAACTCGGTGAAAGCCACACCTCTGTCTTGAATAACA 1314
QY 1142 CTGGCACTGGAACATCTGCTGATGATCTGTTCTCAATATATCAAGGCTTCTGAGTGA 1201
Db 1315 CTGGCACTGGAACATCTGCTGATGATCTGTTCTCAATATATCAAGGCTTCTGAGTGA 1374
QY 1202 CACAAGTAGGCTCCCAATAAAG 1224
Db 1375 CACAAGTAGGCTCCCAATAAAG 1397

RESULT 3
LOCUS BD091714 1652 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel deoxyribonuclease, gene encoding thereof and use thereof.
ACCESSION BD091714
VERSION BD091714.1 GI:22637325
KEYWORDS WO 0112793-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1652)
Tanuma, S. and ShioKawa, D.
Novel deoxyribonuclease, gene encoding thereof and use thereof
Patent: WO 0112793-A 1 22-FEB-2001;
SEIICHI TANUMA,DAISUKE SHIOKAWA
```

```
COMMENT OS Mus musculus (mouse)
PN WO 0112793-A/1
PD 22-FEB-2001
PF 01-MAY-2000 WO 2000JP002893
PR 17-AUG-1999 JP 99P 230870
PT SEIICHI TANUMA,DAISUKE SHIOKAWA
PC C12N9/22,C12N15/12,C12N5/10,C12N7/01,A61K31/195,A61K48/00 CC
FH Key Location/Qualifiers
FT CDS (213)..(1274).

FEATURES
source
1..1652
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 99.7%; Score 1219.8; DB 6; Length 1652;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAGTCCCTCGATGGAATGAGGCCACAGATAGAAAATGACAGCAAAAGCCTCTAAGAA 61
Db 175 CACAGTCCCTCGATGGAATGAGGCCACAGATAGAAAATGACAGCAAAAGCCTCTAAGAA 234
QY 62 CAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
Db 235 CAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 294
QY 122 GCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
Db 295 GCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
QY 182 CTAGCAGGCAAGTGAAGAGGCGGGCTGCACTGATGATGATGATGATGATGATGATGATGAT 241
Db 355 CTAGCAGGCAAGTGAAGAGGCGGGCTGCACTGATGATGATGATGATGATGATGATGATGAT 414
QY 242 CCGTGAACAGAGGCTCTACCTGATTTACAGACACAGAGGATGCTCTGGGGAGGACCTTAC 301
Db 415 CCGTGAACAGAGGCTCTACCTGATTTACAGACACAGAGGATGCTCTGGGGAGGACCTTAC 474
QY 302 AGCATCTGTATGACACACATTAATCCAGAAATGACACAGCCTTATCTAATATATCAAGATG 361
Db 475 AGCATCTGTATGACACACATTAATCCAGAAATGACACAGCCTTATCTAATATATCAAGATG 534
QY 362 GTGTCCCTGGATCTGTGAAATTAACAGACAGATGAGCAATGCCAAAGGTCTGCTGGTAT 421
Db 535 GTGTCCCTGGATCTGTGAAATTAACAGACAGATGAGCAATGCCAAAGGTCTGCTGGTAT 594
QY 422 GGAACAGAACGAGGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATG 481
Db 595 GGAACAGAACGAGGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATG 654
QY 482 GCTATGAGTACCAACCTCGGGGAGGATATGGACAAACCGGCATCTGCACTCACTTCG 541
Db 655 GCTATGAGTACCAACCTCGGGGAGGATATGGACAAACCGGCATCTGCACTCACTTCG 714
QY 542 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 601
Db 715 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 774
QY 602 GCTGCTTCATTCACAGCATTTCATCTGAAACCTTATCTACATGCCCCGAGATGTCGA 661
Db 775 GCTGCTTCATTCACAGCATTTCATCTGAAACCTTATCTACATGCCCCGAGATGTCGA 834
QY 662 ACTCCAGTTTCTTAAAGATCCCTGTGCTGGTACCTGCTGAACTGCACTGAGCCAGGTC 721
Db 835 ACTCCAGTTTCTTAAAGATCCCTGTGCTGGTACCTGCTGAACTGCACTGAGCCAGGTC 894
QY 722 TAAACTTCGTTCCATTTTGCAAAATCAAGTTTATATCTGATGACATCTTTACAGGATGA 781
Db 895 TAAACTTCGTTCCATTTTGCAAAATCAAGTTTATATCTGATGACATCTTTACAGGATGA 954
QY 782 TAGCTCAAAAGTTGAAGACACATTTGTTAGCAAAACCTGGCAGAAAAAGAAACAGAGC 841
```

Db	955	TAGCTCAAAGTTGAAGACACATTTGTTAGCACAACCTGGCAGAAAAAGAAACACAGAGC	1014
Qy	842	TTCCCTCAAACCTGTTCCCTGCCCTTACCATGTCTACAACATCAAGTCCATTGGGGTTAACTT	901
Db	1015	TTCCCTCAAACCTGTTCCCTGCCCTTACCATGTCTACAACATCAAGTCCATTGGGGTTAACTT	1074
Qy	902	CCAGAGCTTACTTCAGTTCTCGCCACAGACCATTCCAAATGGTGTGTTTCCATAAAGGGCT	961
Db	1075	CCAGAGCTTACTTCAGTTCTCGCCACAGACCATTCCAAATGGTGTGTTTCCATAAAGGGCT	1134
Qy	962	CCGCAAAATCGTGGACCTGCATTTGGAGACCTAAATCGAAGCCTACCAAGCCTTTAAGAG	1021
Db	1135	CCGCAAAATCGTGGACCTGCATTTGGAGACCTAAATCGAAGCCTACCAAGCCTTTAAGAG	1194
Qy	1022	GTGAGAGATTCACTGTACAAAGAAATCACTACATTTACCGAGCATTTCAATAATTATATC	1081
Db	1195	GTGAGAGATTCACTGTACAAAGAAATCACTACATTTACCGAGCATTTCAATAATTATATC	1254
Qy	1082	TCCGTTATGGTTCCTGTAAAGTAAACTCGGTGAAGGCGCACACCTCTGTCTCTTCAAAAACA	1141
Db	1255	TCCGTTATGGTTCCTGTAAAGTAAACTCGGTGAAGGCGCACACCTCTGTCTCTTCAAAAACA	1314
Qy	1142	CTGGCACTGGAAACATCTCGCCTTGGATCTGTTTCCATAATATCAAGCCTTCTGAGTGAG	1201
Db	1315	CTGGCACTGGAAACATCTCGCCTTGGATCTGTTTCCATAATATCAAGCCTTCTGAGTGAG	1374
Qy	1202	CACAACGTAGCGTCCATAAAG	1224
Db	1375	CACAACGTAGCGTCCATAAAG	1397
RESULT 4			
AF128888		1654 bp	linear
LOCUS		Mus musculus deoxyribonuclease D1AD mRNA	complete cds.
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			


```

Db      1135  CCGCAATCGCTGGACCTGCGATTCGAGACCTTAAATCGAAGCCTACACCAAGCCTTAAGAG 1194
Qy      1022  GTGAGGATTCATCTGTACAAAGAAATCACTACATTTACCGGCAATTCATAAAATATATC 1081
Db      1195  GTGAGGATTCATCTGTACAAAGAAATCACTACATTTACCGGCAATTCATAAAATATATC 1254
Qy      1082  TCGGTTATGGTTCCTGTAAAGTAAACTCGGTGAAAGGCCACACCCCTCTGTCTCTGAAACA 1141
Db      1255  TCGGTTATGGTTCCTGTAAAGTAAACTCGGTGAAAGGCCACACCCCTCTGTCTCTGAAACA 1314
Qy      1142  CTGCACTGGAACATCTCGCTTGAGATCTCTTCCATAATATCAAGGCTTCTGAGTGA 1201
Db      1315  CTGCACTGGAACATCTCGCTTGAGATCTCTTCCATAATATTCAGGCTTCTGAGTGA 1374
Qy      1202  CACAACGTAGCTCCCAATAAAG 1224
Db      1375  CACAACGTAGCTCCCAATAAAG 1397

RESULT 5
AF178974
LOCUS      1498 bp mRNA linear ROD 29-NOV-1999
DEFINITION Rattus norvegicus deoxyribonuclease DIAD mRNA, complete cds.
ACCESSION AF178974
VERSION AF178974.1 GI:6470130
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1498)
          Tanuma,S. and Shiokawa,D.
          Cloning of a cDNA encoding a rat DNase II-like acid DNase
          Biochem. Biophys. Res. Commun. 265 (2), 395-399 (1999)
          20025354
          PUBMED 10558878
REFERENCE 2 (bases 1 to 1498)
          Shiokawa,D. and Tanuma,S.
          Direct Submission
          Submitted (19-AUG-1999) Biochemistry, Science University of Tokyo,
          Shinjuku-ku Ichigaya Funagawara-machi 12, Tokyo 162-0826, Japan
FEATURES
          source
          1..1498
          Location/Qualifiers
          /organism="Rattus norvegicus"
          /mol_type="mRNA"
          /strain="Wistar"
          /db_xref="taxon:10116"
          159..1229
          /codon_start=1
          /product="deoxyribonuclease DIAD"
          /protein_id="AA113596.1"
          /db_xref="GI:6470131"
          /translation="MTAQLKALPILFVALSGVLGTPVISINEDGKAVDWFAFYKL
          PARTSRGGTGLDLYLSDNRTWSKSHLINSRSLRGTLEOLYRAHNKNDTAY
          LYNDAPASVYNSYNGYHAKGLLVNVRVQGFLLHSIPKPPVEKGYEVPSSROY
          AQSGLCTIKTSQYETIDSLVFPNTVYSCFIPNIFRWELHMPQMAKSKSLPS
          RLTLVLSAOGNLFHFAKSTFTYDDIIFAWIAQKLKVLHLYVESQKNNHLPSCSL
          PLYVNIYKALRGPLSDPFSHDHDKWCVSTKDSQARWTCIGDLNRSFQALRSGGFI
          CSKNRYIIQSFDRFLVSHYASCN"

CDS
          70.9%; Score 867.6; DB 10; Length 1498;
          Best Local Similarity 82.8%; Pred. No. 8.6e-249;
          Matches 1017; Conservative 0; Mismatches 204; Indels 7; Gaps 2;

Qy      2    CCCAGTCCCTGCATGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTTAAGAA 61
Db      121   CACAGTCTGAGCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTTAAGAA 180
Qy      62   CAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
Db      181   CAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

```

```

Qy      122   GCAGAAATGATATATGTTGAAGCTGTGGAATCTGGTTTATCTTTTAAAGTACCCAAAAGGA 181
Db      241   GCATAAAGCAAGATGTTAAAGCCGTGACCTGGTTTGCCTTTTATAAGTTACCCAGAAGGA 300
Qy      182   CTAGCAGGCAAGTGAAGAGCGCGGCTGCAGTACCTGTACCTGTGAGCTCCACACAGACAA 241
Db      301   CAGCAGAGAGGTTACAGGGATGGGGCTGGATTACCTGTACCTGTGAGCTCCACAAATGAGAA 360
Qy      242   CCTGGAAACAAGAGCCTTACCTGATTAAACAGCACAGGAGTGTCTCTGGGAGGACCTTAC 301
Db      361   CCTGGACCAAGAGCCACCACCTGATTAAACAGCAGCAGAGAGCTCCCTGGGAGAGACCTGG 420
Qy      302   AGCATCTGTATGACACACATAATTCACAGATAGACACAGCAGCCTATCTTAATATACAAAGATG 361
Db      421   AGCAGCTGTATGAAGCACACAATGCCAAGATGACACAGCCTATCTGATATACAATGAGC 480
Qy      362   GTGTCCCTGGATCTGTGAATTTACAGCAGACAGATATGACATGCCAAGGTCTGCTGTAT 421
Db      481   CTGTCCCTGGATCTGTGAATTTACAGCAGAAATTAACGGGCATGCCAAGGTCTGCTGTAT 540
Qy      422   GGAACAGAACCGCAGGGGTTCTGGCTGATACACTCTGTTCCTCAAGTTTCCCCAGTTTCC 478
Db      541   GGAACAGAGTGCAGGGGTTCTGGCTGATTCATTCTATTCCTCAAGTTTCCCCAGTTTCCGG 600
Qy      479   ---ATGGCTATGATACCCAACTCCGGGAGGCGATATGACAAACCGGCATCTGCATCA 535
Db      601   AAAAAGGCTATGAATACCCAAAGCTCGGGAGGCAATATGACAAAGTGGCTCTGCATCA 660
Qy      536   CTTTCCGATACAGCCAGTTTGAAGAAATAGATTTTCAGCTCTTGGTCTTTCACACCAACA 595
Db      661   CTCTCAATACAGCCAGTATGACAAATAGATTTCTCAGCTCTTGGTCTTTCACGCCAAACA 720
Qy      596   TCTACAGCTGCTTCATTCCAAGCACCTTTCACTGGAACCTTATCTACATGATGATGATGATG 655
Db      721   TCTACAGCTGTTTCAATCCCAACATCTTTCGCTGGGAATCTCATCCACATGATGATGATG 780
Qy      656   GTGCCAACTCCAGTCTCTTAAAGATCCCTGTGCGGATCCCTGTGAACTGCACTCAGCCC 715
Db      781   GTGCCAAGTCCAGTGCCTCAAAAGATCCCTAGCGCGGCTCACTGTACTTCACTCAGTCCAG 840
Qy      716   AGGCTCTAAACTTCGTCATTTTGAAATCAAGTTTATCTATCTGATGATGATGATGATGATG 775
Db      841   AGGCTCTAAACTTCCTCAATTTTGAAATCAACTTTTATCTATCTGATGATGATGATGATG 900
Qy      776   GATGATAGCTCAAAAGTTGAAGACACATTTGTAGACAAACCTGGCAGCAAAAGAAAC 835
Db      901   CCTGGATAGCTCAAAAGCTGAAGGTGCATTTCTGTGTAGAACTCTGGCAGCGAAGAAC 960
Qy      836   AAGAGCTTCTTCAACATGTTCCCTGCTTACCATGTCTACAAATCAAGTCCATTTGGGG 895
Db      961   ACGAGCTTCTTCAAACTGTTCCCTGCTTACCATGTCTACAAATCAAGGCAATTTGGGG 1020
Qy      896   TAACTTCCAAAGTCTTACTTCACTTCCGCAAGACCATTCGAAATGGTGTGTTCCATTA 955
Db      1021   GACCTTGCAGTCCGACTTCCCTTCTCATCAGACCATTCGAAATGGTGTGTTTCCACAA 1080
Qy      956   AGGGCTCCGCAAAATCGCTGGACCTGCATTGGAGACCTAAATCGAAGCCTTACCAAGCCT 1015
Db      1081   AGGACTCTCAGGCTCGCTGGACCTGTATTGGAGACTTAATCGGAGCCCATCAAGCCT 1140
Qy      1016   TAAGAGTGGAGGATTCATCTGTACAAAGATCACTACATTTACAGGCAATTTCAAT 1075
Db      1141   TGAGAAGTGGAGGATTTATCTGTTCGAAGAATCGGTACATTTTACCAATCATTTGATAGT 1200
Qy      1076   TATATCTCCGTTATGGGTTCTGTAGTAACTCGGTGAAGGCCACACCTCTGCTCTG 1135
Db      1201   TAGTTTCCATATAGCTCTCTGTACTAAGCTTGGGGAAGGTCCCAACCTCTGCCCTTA 1260
Qy      1136   AAAACACTGGCACTGGAAACATCTCGCTTGGATCTGTTCTCATATATCAAGGCTCTG 1195
Db      1261   AAGACACTGGCTCTGGGACATCTTGGCTTGGATCTATTTCTCATATGTTAAAGTTTCTG 1320

```

```
QY 1196 AGTGACCAACCTAGCTCCATATAA 1223
Db 1321 AGTG-TCCTAACATCTCATCCATAAAA 1347

RESULT 6
AR201144
LOCUS AR201144 1268 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6358723.
ACCESSION AR201144
VERSION AR201144.1 GI:20252032
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1268)
AUTHORS Eastman,A.Richard, and Krieser,R.Joe.
TITLE Deoxyribonuclease II.beta. proteins and cdnas
JOURNAL Patent: US 6358723-A 3 19-MAR-2002;
FEATURES
    Location/Qualifiers
        1..1268
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
Query Match 55.0%; Score 672.8; DB 6; Length 1268;
Best local Similarity 75.0%; Pred. No. 2.6e-190;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 31 AGATAGAAATGACAGCAAAAGCCCTAAGAACAGTCTCTTTCTGCTCTCTTTGGCCCTC 90
Db 37 AAACAGAAATGATGGCAAGACTCTGTAAGACATCTTTGCTTTGCTCTCTTTGGCCCTC 96
QY 91 TCTGGGTCTCTGGGACACACCAAGAAATCTCATGCGAAGATGAATATGTTGAAGCTGTGGAC 150
Db 97 TTTGGGTGCTGGGGGAGCAACAATTCATGCGAAGATGAAGAAAGGAAAGCTGTGGAC 156
QY 151 TGGTTATCTTTTATAGTTACCAAGAGGACTAGCAAGCAAGTGAAGCGCGGGCTG 210
Db 157 TGGTTATCTTTTATAGTTACCAAGAGGACTAGCAAGCAAGTGAAGCGCGGGCTG 216
QY 211 CAGTACCTGTACCTGGACTCCACAGCAAAACCTCGGAACAGAGCCTCTACCTGATTAAAC 270
Db 217 GAGTACCTGTACCTAGACTCTACACTAGAGCTGGAGGAAGAGTGAGCAACTAATGAAT 276
QY 271 AGCACCGAGAGTCTCTGGGAGGAGCCTTACAGCATCTGTATGACACATATAATCCACG 330
Db 277 GACACCAAGAGTGTCTGGGAAGGACATTAACAAGCTATATGAAGCATATGCCCTTAAG 336
QY 331 AATG---ACAGAGCTATCTAATATACAAAGATGGTCCCTGGATCTGTGAATTTACAGC 387
Db 337 AGTAACAACACAGGCTATCTAATATACAAATGATGGAGTCCCTTAACCTGTGAATTTACAGT 396
QY 388 AGACAGTATGACATGCCAAAGGCTGCTGTATGGAACAGAACAGAGGTTCTGGCTG 447
Db 397 AGAAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456
QY 448 ATACACTGTGTCCAAAGTTTCCCCCAAGTTC-----ATGGGTATGAGTACCCCACTCG 501
Db 457 ATTCAATCCATCCCTCAGTTTCCCTCCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516
QY 502 GGGAGGCATATGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTCAGGAA 561
Db 517 GGGAGCAAAATGGCAAGAGTGGCAATCTGCATCACTTTCAAGTACACACAGTATAGGCA 576
QY 562 ATAGATTTTTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTGTTTCATTCACAGACC 621
Db 577 ATAGATTTTCAGCTCTTGGTCTGCAACCCCAAGCTATAGCTGTCCATCCAGCCACC 636
QY 622 TTTCTGTGAATCTTATCTATGATGCCCGGATGTGTGCCAATCCAGTTCCTTAAGATC 681
Db 637 TTTTCCAGGAGCTCATTCATATGCCCGGAGCTGTGCCAGCGGGCCAGCTCATCAGAGATT 696
```

```
QY 682 CTTGTCGGTACCTCGCTGAAGCTGCACTCAGCCAGGGTCTAAACTTGTGCAATTTTGA 741
Db 697 CTTGGCAGGCTCTCTACCACTCTTCACTAGTCGGGACAGGACAAATAATCTCCATTTTGA 756
QY 742 AATCAAGTTTATATCTGATGACATCTTTACAGGATGGATAGCTCAAAAGTTTGAAGACA 801
Db 757 AAGTCGGAATCTTTCTTGAAGACATCTTTGACGCTTGGATGGCTCAACGGCTGAGACA 816
QY 802 CATTCTGTAGCACAACCTGGCAGAAAAAAGAAACAAAGAGCTTCTCTTCAAACTGTTCCTG 861
Db 817 CACTTGTATACAGAAACCTGGCAGCGAAGAAAGACAGAGCTTCTCTTCAAACTGTTCCTG 876
QY 862 CTTTACCATGTCTACAAACATCAAGTCAATTTGGGGTAACTTCCCAAGCTTCACTTCACTTCT 921
Db 877 CTTTACCATGTCTACAAATATAAAAGCAATTTAAATTTATCAGCACATCTTTATTTTCACTTCT 936
QY 922 CGCAGACACATTTCCAAATGCTGTGTTTCCATAAGGGCTCCGCAATCGCTGACCTGC 981
Db 937 TATCAAGATCACCGCAAGTGGTGTATTTTCCAAAGGGACCAAAATCGCTGGACATGT 996
QY 982 ATTGGAGACCTTAATCGAAGCTTACACCAAGCCTTAAGAGGTGAGAGTCAATCTGTACA 1041
Db 997 ATTGGAGACCTTAATCGGAGTCCACCAAGCCTTCAGAAGTGGAGGATTCATTTGTACC 1056
QY 1042 AAGAATCATCATTTTACCAAGGATTTTCAAAATATATCTCGTTATGGGTTCTGTAAAG 1101
Db 1057 CAGAAATGGCAATTTTACCAAGCATTTCAAGGATTTAGTATTATATCTATGAAGCTGTAAAG 1116
QY 1102 TAACTCGGTGAAGGCCACACACCTCTGCTCTTGAACACTGGCAGCTGGAAACATCTCGC 1161
Db 1117 TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAACCTTGACAAATGGGTCTTCTTCC 1176
QY 1162 CTTGGATCTGTTCTCCATAATA 1183
Db 1177 ATTACACCTTTTATATTTTA 1198

RESULT 7
AF274571 1140 bp mRNA linear PRI 12-JUN-2001
LOCUS AF274571 Homo sapiens deoxyribonuclease II beta mRNA, complete cds.
DEFINITION AF274571.1 GI:8571463
ACCESSION AF274571.1
VERSION AF274571.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1140)
AUTHORS Krieser,R.J., MacLea,K.S., Park,J.P. and Eastman,A.
TITLE The cloning, genomic structure, localization, and expression of human deoxyribonuclease IIbeta
JOURNAL Gene 269 (1-2), 205-216 (2001)
MEDLINE 21272509
PUBMED 11376952
REFERENCE 2 (bases 1 to 1140)
AUTHORS Krieser,R.J. and Eastman,A.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-2000) Pharmacology, Dartmouth Medical School, 7650 Remsen, Hanover, NH 03755, USA
FEATURES
    Location/Qualifiers
        1..1140
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /chromosome="1"
        /map="1p22.3"
        /note="additional 3'-untranslated sequence up to a potential polyadenylation site is present in several EST clones in the database"
        46..1119
        /note="acid endonuclease"
        /codon_start=1
CDS
```

```
/product="deoxyribonuclease II beta"
/protein_id="AAF76893.1"
/db_xref="GI:8571464"
/translation="MMARLRTSPALLFLGLFGLGAATISCRNEKXAVDWTFYKL
PKQKNGSGTGLYLDSTTSWRKSEQLMNDKLSVIGRTIQOLLYEAVASKNNTA
YLIYNDVFPKPNYSRIKHTKGLLLNVRVQGWFLHISIFQFPPIPEEGDYPTGRR
NGSGGICITPKYQYEAIDSLVPCNPVYSCSIPATFHQELIHPOLCRASSEIP
GRLLTTLQSAQGGKFLHFAKSDSFLDDIFAAWMAQRILKTHLLTETWQKEQLPSNCS
LPYHVNKIKALSRHSYFSSYODHAKWCIQSQTKRWTICIGDLNRSHPQAFRSQGF
ICTQNMQIYQAFGLVLYESCK"
ORIGIN
Query Match 53.4%; Score 653.4; DB 9; Length 1140;
Best Local Similarity 75.9%; Pred. No. 1.8e-184;
Matches 836; Conservative 0; Mismatches 256; Indels 9; Gaps 2;
QY 31 AGATAGAAAATGACAGAAAGCTCTTAAGAACAGTCTTCTTTTGTCTCTTTTGCCTC 90
DB 37 AAACAGAAAATGATGGCAAGACTGCTAAGAACATCTCTTTTGTCTCTTTTGCCTC 96
QY 91 TCTGGGCTCTCTGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
DB 97 TTTGGGCTGTGGGACACAGCAATTTCAATGAGAAATGAGAGGAAAGCTGTGGAC 156
QY 151 TGGTTATCTTTTATAGTTACCAAGACTAGCAAGCAAGTGAAGAGGGGGGGCTG 210
DB 157 TGGTTTACTTTTATAAGTTTACCTAAAGACAAACAAAGAAAGTGGAGAGCTGGGTTA 216
QY 211 CAGTACTGTACTGGACTCCACAGACAACTGGAACAGAGCCCTCTACCTGATTAAC 270
DB 217 GAGTACTGTACTGACTCTCACTAGAGCTGGAGGAGAGTGAGCACTAAATGAAT 276
QY 271 AGCACCAGAGTCTCTGGGAGGACTTACAGCATCTGTATCACACACATATTCACG 330
DB 277 GACACCAAGAGTGTTTGGAGAGCACTTACACAGCTATATGAGCAATGCTCTAAG 336
QY 331 ---AATGACAGCCCTATCTAATATATCAACGATGGTGTCCCTGGATCTGTGAATTACAGC 387
DB 337 AGTAAACAACAGACGCTATCTAATATATCAATGATGGAGTCCCTAAACCTGTGAATTACAGT 396
QY 388 AGACAGTANGGACATGCCAAGTCTGCTGGTATGAAACAGACGAGGGGTCTGGCTG 447
DB 397 AGAAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456
QY 448 ATACACTCTGTCTCCCAAGTTTCCCCAGTTC-----ATGGCTATGATGACCAACCTCG 501
DB 457 ATTCAITTCATCCCTCAGTTTCTCCAAATCCGGAAGAGGCTATGATATCCACCACA 516
QY 502 GGGAGCGGATATGGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGAAGAA 561
DB 517 GGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACAGTATGAGCA 576
QY 562 ATAGATTTTCAGCTCTGTGCTTACACCAACAATCTACAGCTGCTTATTCGACGACC 621
DB 577 ATAGATTTTCAGCTCTGTGCTGTCGACCCCAACGCTATAGCTGCTCCATCCAGCCACC 636
QY 622 TTTCACTGGAACCTTATCTACATGCCCCGATGTGCGCAACTCCAGTTTCCTTAAGATC 681
DB 637 TTTCAACGAGGCTCATTCAGATGCCAGCTGTGACCCGAGGCGACGCTCATCAGAGAT 696
QY 682 CTTGCGGTAACCTGCTGAACTGCACTCAGCCCCAGGGTCTAAACTCTGTCATTTTGA 741
DB 697 CTTGCGGCTCCTCACCACTTCACTGTCGCCCGGACAAAATTCCTCCATTTTGA 756
QY 742 AAATCAAGTTTATATCTATGATCATCTTTACAGATGATAGCTCAAAAGTTGAAGACA 801
DB 757 AAGTCGGAATCTTTCTTGACACATCTTTGACGCTGGATGGCTCAACGGCTGAAGACA 816
QY 802 CATTTGTTAGCAAAACCTGGCAGAAAAGAAAGAAAGAGCTTCCCTTCAACCTGTTCCCTG 861
DB 817 CACTTGTAAACAAACCTGGCAGGAAAGAAAGAGAGCTTCCCTTCAAACTGCTCCCTT 876
QY 862 CTTTACCAGTCTACAACTCAAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCACTTCT 921
```

```
DB 877 CCTTACCATGTCATCAATATAAAGCAATTAAATATACAGCACTCTTATTTCACTTCT 936
QY 922 CGCCAGACCAATTCCAAATGGTGTGTTTCCATAAAGGGCTCCGCAAAATCGCTGGACCTGC 981
DB 937 TATCAAGATCAGCGCAAGTGGTGTATTTCCAAAAGGGCACCAAAATCGCTGACATGT 996
QY 982 ATTGGAGACCTAATTCGAAGCCTACCAAGCCTTAAGAGTGGAGATTCAATCTGTACA 1041
DB 997 ATTGGAGACCTAATTCGAAGCCTTCCACCAAGCCTTCCAGAAAGTGGAGATTCAATTTGTA 1056
QY 1042 AAGAATCACTACATTTTACCAGGCAATTTATAAATATATATCTCCGTTATGGTCTGTAA 1101
DB 1057 CAGATTTGGCAATTTTACCAGGCAATTTCAAGGATTAGTATATATACTAAGAAAGCTGTA 1116
QY 1102 TAAACTCGGTGAAGGCCACA 1122
DB 1117 TAAACTTGGTGAAGGACACA 1137
RESULT 8
AR432047 1086 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 4 from patent US 6653118.
ACCESSION AR432047
VERSION AR432047.1 GI:40194250
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1086)
AUTHORS Tanuma,S.-i. and Shiohawa,D.
TITLE Deoxyribonuclease, Gene encoding same and use thereof
JOURNAL Patent: US 6653118-A 4 25-NOV-2003;
FEATURES Location/Qualifiers
source
1..1086
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 52.4%; Score 641.8; DB 6; Length 1086;
Best Local Similarity 75.9%; Pred. No. 5.4e-181;
Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;
QY 31 AGATAGAAAATGACAGAAAGCTCTTAAGAACAGTGTCTTTTGTCTCTTTTGTCTCTT 90
DB 4 AAACAGAAAATGATGGCAAGACTGTGTAAGAACATCTTTTGTCTCTTTTGTCTCTT 63
QY 91 TCTGGGCTCTCTGGGACACCAAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
DB 64 TTTGGGCTCTGGGCGACGACAAATTTTCATGAGAAATGAAGAGGAAAGCTGTGGAC 123
QY 151 TGGTTTATCTTTTATAAGTTTACCAGGCAAGTGTAGCAAGGCAAGTGAAGCGGGGCTG 210
DB 124 TGGTTTACTTTTATAAGTTTACTTAAAGACAAACACAGAAAGTGGAGAGACTGGGTTA 183
QY 211 CAGTACTGTACTGGACTTCCAAAGACAAACCTGGAACAGAGCTCTTACCTGATTAAC 270
DB 184 GAGTACTGTACTAGACTCTCAACTAGAGCTGGAGGAGAGTGAAGCACTAATGAAT 243
QY 271 AGCACCAGAGGCTCTCTGGGAGGAGCTTACAGCATCTGTATGACACACATATTTCCAG 330
DB 244 GACACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATGAAGCATATGCTCTAAG 303
QY 331 ---AATGACAGCCCTATCTAATAACAGATGTGTCCCTGGATCTGTGAATTACAGC 387
DB 304 AGTAAACAACAGCTTATCTAATATATCAATGATGAGTCCCTTAAACCTGTGAATTACAGC 363
QY 388 AGACAGTATGACATGCGCAAGGTTCTGCTGATGAAACAGAAACGAGGGGTTCTGGCTG 447
DB 364 AGAAAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTG 423
QY 448 ATACACTCTGTCTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTCG 501
```

Db	424	ATTCAATCCATCCCTCAGTTTCTCCATTCGGAAGAGCTATGATATCCACCCACA	483
Qy	502	GGGAGGCGATATGACAAACCGGCATCTGATCACTTCGGAATACAGCCAGTTTGAGGAA	561
Db	484	GGGAGAGGAATGGACAAAGTGGCATCTGATTAACCTTCAAGTACACCAAGTATGAGCA	543
Qy	562	ATGATATTTTCACTTCTGTTCTTACACCAACATCTACAGCTGCTTCATTCACAGCACC	621
Db	544	ATGATATTTTCACTTCTGTTCTTACACCAACATCTACAGCTGCTTCATTCACAGCACC	603
Qy	622	TTTCACTGGAACTTATCTACATGCTCCCGGATGCTGCTCAACTCCAGTTCCTTAAAGATC	681
Db	604	TTTCACTGGAACTTATCTACATGCTCCCGGATGCTGCTCAACTCCAGTTCCTTAAAGATC	663
Qy	682	CTGTCCGGTACCTCGTGAATGCACTCAGCCAGGCTTAAATCTTCGTCCTTAAAGATC	741
Db	664	CTGTCCGGTACCTCGTGAATGCACTCAGCCAGGCTTAAATCTTCGTCCTTAAAGATC	723
Qy	742	AAATCAAGTTTATATCTGATGATCTTTACAGATGATGCTCAAAAGTTCGAAGACA	801
Db	724	AAATCAAGTTTATATCTGATGATCTTTACAGATGATGCTCAAAAGTTCGAAGACA	783
Qy	802	CATTGTGTAGCACAACTCTGACAGAAAGAAACAAAGAGCTTCTTCAAACTGTTCCCTG	861
Db	784	CATTGTGTAGCACAACTCTGACAGAAAGAAACAAAGAGCTTCTTCAAACTGTTCCCTG	843
Qy	862	CCTTACCATGCTACACATCAAGTCCATGCTGGGTAACTTCCAACTGTTCACTGTTCT	921
Db	844	CCTTACCATGCTACACATCAAGTCCATGCTGGGTAACTTCCAACTGTTCACTGTTCT	903
Qy	922	CGCAAGACCATTCCTAAATGCTGTTTCCATTAAGGCTCCGCAATCGCTGGACCTGC	981
Db	904	TATCAAGATCATGCGCAAGTGTGTATTTCCCAAGAGGCAACAAATTCGCTGGACATGT	963
Qy	982	ATTGAGACCTTAAATCGAAGCTTACCAAGCTTAAAGAGTGGAGGATTCATCTGTACA	1041
Db	964	ATTGAGACCTTAAATCGAAGCTTACCAAGCTTAAAGAGTGGAGGATTCATCTGTACA	1023
Qy	1042	AAGAATCATCTACATTTACAGGATTTTCAATTAATATCTCGTATGTTGTTCTGTAAAG	1101
Db	1024	CAGAATTTGGCAATTTTACCAAGATTTTCAAGGATTTAGTATATATGATGAAAGCTGTAAG	1083
Qy	1102	TAA 1104	
Db	1084	TAA 1086	

RESULT 9
BD091715
LOCUS BD091715
DEFINITION Novel deoxyribonuclease, gene encoding thereof and use thereof.
ACCESSION BD091715
VERSION BD091715.1 GI:22637326
KEYWORDS WO 0112793-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1086)
Tanuma, S. and Shikawa, D.
Novel deoxyribonuclease, gene encoding thereof and use thereof
Patent: WO 0112793-A 2 22-FEB-2001;
SEIICHI TANUMA, DAISUKE SHIKAWA
OS Homo sapiens (human)
PN WO 0112793-A/2
PD 22-FEB-2001
PF 01-MAY-2000 WO 2000JP02893
PR 17-AUG-1999 JP 99P 230870
PI SEIICHI TANUMA, DAISUKE SHIKAWA
PC C12N15/12, C12N15/10, C12N7/01, A61K31/195, A61K48/00 CC
FH Key Location/Qualifiers
FT CDS (1)..(1083).

FEATURES	Location/Qualifiers
source	1..1086
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
	Query Match 52.4%; Score 641.8; DB 6; Length 1086;
	Best Local Similarity 75.9%; Pred. No. 5.4e-181;
	Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;
Qy	31 AGATAGAAATGACAGCAAAAGCCTTAAGAACAGTCTTTCTTTGCTCTTTTGGCCCTC 90
Db	4 AAACAGAAATGATGGCAGACACTGCTAAGAACATCCCTTTGCTTTCTTCTTTGGCCCTC 63
Qy	91 TCTGGGCTCTGGGGACACAGAAATCTCATGACAGAAATGAATATGCTGAAGCTGTGGAC 150
Db	64 TTTGGGCTCTGGGGGAGCAACAATTTCAAGCAAAATGAAGAGGAAAGCTGTGGAC 123
Qy	151 TGGTTTATCTTTTATAGTTTACCCAAAGGACTAGCAAGCAAGTGAAGAGGCGGGCTG 210
Db	124 TGGTTTATCTTTTATAGTTTACCTTAAAGACAAAACAAGGAAAGTGGAGAGCTGGGTTA 183
Qy	211 CAGTACCTGTACCTGGGCTCAGACAGCAAACTGGGAACAGAGCCTCTTACCTGATTAAC 270
Db	184 GAGTACCTGTACCTGAGCTCTCAACTAGAGCTGGAGGAGAGTGAAGCACTAATGAAT 243
Qy	271 AGCACCAGGAGTGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATAATTTCCACG 330
Db	244 GACACCAGGAGTGTCTTGGGAGGACATTAACACAGCTATATGAAGCATATGCTCTTAAG 303
Qy	331 ---AATGACACAGCCTATCTAATAATACAAAGATGCTGCTTGGATCTGTGAATTAACG 387
Db	304 AGTAAACACACAGCCTATCTAATATACATGATGAGTCCCTTAAACCTGTGAATTAACG 363
Qy	388 AGACAGTATGACATGCAAAAGGCTCTGCTGTATGGAACAGAAACGAGGGTCTGGCTG 447
Db	364 AGAAAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTCTTGGCTG 423
Qy	448 ATACACTCTGTTTCCCAAGTTTCCCCAGTTTC-----ATGGCTATGAGTACCCCACTCG 501
Db	424 ATTCACTCCATCTCTCAGTTTCTTCCATTCGGAGAGAGCTATGATATCCACCCACA 483
Qy	502 GGGAGGCGATATGACAAACCGGCATCTGCACTCTTTCGGATACAGCAGCTTTGAGGAA 561
Db	484 GGGAGCGAATGACAAAGTGGCATCTGCACTCTTCAAGTACACACAGTATGAGCA 543
Qy	562 ATAGATTTTCACTCTTGGTCTTACACCAACATCTACAGCTGCTTTCATTCACAGCACC 621
Db	544 ATAGATTTTCACTCTTGGTCTTCAACCCCAAGCTTATAGTCTTCCATCCAGCACC 603
Qy	622 TTTCACTGGAACTTATCTACATGCTCCCGGATGCTGCAACTCCAGTTCCTTAAAGATC 681
Db	604 TTTCACTGGAGCTCAATTCATGCTCCCGGATGCTGCAACTCCAGGCTTCACTCAGAGAT 663
Qy	682 CCGTCTCGGTACCTCGTGAATGCACTCAGCCAGGCTTAAATCTTGTGCTCAATTTTGA 741
Db	664 CCGTCTCGGTCTCTTCAACACACTTCACTGCGGCCCGGAGGACAAAATTCCTTCAATTTGA 723
Qy	742 AAATCAAGTTTATCTGATGACATCTTACAGGATGATGATGATGATGATGATGATGATG 801
Db	724 AAGTCGGATTTCTTTTCTTGTATGGCACTTTTGCAGCTGATGCTTCAACGGCTGAAGCA 783
Qy	802 CATTGTGTAGCACAACTCTGGCAGAAAAGAAACAAAGAGCTTCTTCAAACTGTTCCCTG 861
Db	784 CATTGTGTAGCACAACTCTGGCAGAAAAGAAACAAAGAGCTTCTTCAAACTGTTCCCTG 843
Qy	862 CTTTACCATGCTACACATCAAGTCCATGCTGGGTAACTTCCAACTGTTCACTGTTCT 921
Db	844 CTTTACCATGCTACACATCAAGTCCATGCTGGGTAACTTCCAACTGTTCACTGTTCT 903
Qy	922 CCGCAGAGCCTTCCAAATGCTGTTTCCATAAAGGGCTCCGCAAAATCGCTGGACCTGC 981

```

Db      904  TATCAAGATCATGCAAGTGGTGTATTTCCAAAAGGCGCACAAAATCGTGGACATGT 963
Qy      982  ATTGGAGACCTAAATCGAAGCTTACACCAAGCCTTAAGAGGTGGAGATTCATCTGTACA 1041
Db      964  ATTGGAGACCTAAATCGAGTCCACCAAGCCTTCAGAAGTGGAGATTCATTTGTACC 1023
Qy      1042  AGAATCACTACATTTACCAGGCAATTCATAAATATATATCTCCGTTATGGTTCCTGTAAG 1101
Db      1024  CAGAAATGGCAAAATTTACCAAGCATTTCAAGGATTAGTATATATATGAAAGCTGTAA 1083
Qy      1102  TAA 1104
Db      1084  TAA 1086

RESULT 10
AF333389
LOCUS   Homo sapiens lung-specific DLAD mRNA, complete cds. PRI 22-NOV-2001
DEFINITION
ACCESSION AF333389
VERSION   AF333389.1 GI:17046270
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1029)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Isolation and characterization of the DLAD/Dlad genes, which lie
          head-to-head with the genes for urate oxidase
JOURNAL   Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)
MEDLINE   21556924
PUBMED    11700027
REFERENCE 2 (bases 1 to 1029)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Direct Submission
JOURNAL   Submitted (04-JAN-2001) Biochemistry, Science Univ. of Tokyo,
          Ichigaya, Shinjuku-Ku, Tokyo 162-0826, Japan
FEATURES
          Location/Qualifiers
            1..1029
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              408..869
              /note="expressed only in lung; does not include first two
              exons"
              /codon_start=1
              /product="lung-specific DLAD"
              /protein_id="AAL34448.1"
              /db_xref="GI:17046271"
              /translation="MPQLCTRASSEIPIGRLLTTLQSAOGOKFLHFAKSDSFLDDIFA
              AWAQRLKTHLLTETWKRQRLPSNCSLPYHVYNIKALKLSHYSYFSYQDHAKWCI
              SQKGTKNRWTCIGDLNRSHPQAFRSGGFICTQNWQIYQAPQGLVLYESCK"

ORIGIN
Query Match 41.8%; Score 512.2; DB 9; Length 1029;
Best Local Similarity 75.7%; Pred. No. 4.4e-142;
Matches 650; Conservative 0; Mismatches 203; Indels 6; Gaps 1;

Qy      331  AATGACACAGCCTATCTAATATACACGATGTGTCCCTGGATCTGTGAATTACAGCAGA 390
Db      90  AACAAACACAGCCTATCTAATATACAAATGATGAGTCCCTTAAACCTGTGAATTACAGCAGA 149
Qy      391  CAGTATGGACATGCCAAAGTCTGTGGTATGGACAGACCGAGGCTCTGGCTGATA 450
Db      150  AAGTATGGACACACCAAGAGTTACTGTGTGGAAACAGAGTTCAAGGGTCTGGCTGATT 209
Qy      451  CACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGATACCCCACTCGGG 504
Db      210  CATTCCATCCCTCAGTTTCTCTCCATTCGGAGAGAGGCTAATGATATCCACCCACAGGG 269
Qy      505  AGGCGATATGGACAAACCGGCATCTGCATCATTTCCGATACAGCCAGTTTGAGGAATA 564

```

```

Db      270  AGACGAATGGAACAAAGTGGCATCTGCATACATTTCAAGTACACACAGTATGAGGCAATA 329
Qy      565  GATTTTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCAATTCGAAGCACCTTT 624
Db      330  GATTTCTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGCTCCATCCAGCCACCTTT 389
Qy      625  CACTGGAAACTTATCTACATGCGCCGAGTGTGCGCAACTCCAGTTCTTAAAGATCCCT 684
Db      390  CACAGAGCTCANTTACATGCCCCCAGCTGTGACCCAGGGCCAGCTCATCAGAGATTCT 449
Qy      685  GTCGGTACCTCGCTGAACCTGCACCTCAGCCCGAGGTCTAAACTTCGTCCATTTTGCAAAA 744
Db      450  GCGAGGCTCTCACCACACTTCAGTCGCGCCAGGACAAAATTCCTCCATTTTGCAAG 509
Qy      745  TCAAGTTTTTATCTGATGACATCTTTTACAGGATGATAGCTCAAAAGTTTGAAGACACAT 804
Db      510  TCGGATTTCTTTTCTTGACGACATCTTTGACGCTGGATGGCTCAACGGCTGAAGACACAC 569
Qy      805  TTCTTAGCACAAACCTGGCAGAAAAGAAACAGAGCTTCCTTCAAACCTGTTCCCTGCCCT 864
Db      570  TTGTTAACGAAACCTGGCAGCGCAAGAGCTTCCTTCAAACCTGCTCCCTCCCT 629
Qy      865  TACCATGTCTACAAATCAAGTCCATTTGGGGTAACTTCCAAAGCTTACTTCAGTTCTCGC 924
Db      630  TACCATGTCTACAAATATAAAGCAATTAATATACAGACACTCTTATTTCACTTTTAT 689
Qy      925  CAAGACATTTCCAAATGGTGTGTTTCCATAAAGGGTCCGCAAAATCGCTGGACCTGCATT 984
Db      690  CRAAGTATGATCCCAAGTGTGTATTTCCCAAAAGGGCACCAAAAATCGCTGGACATGTATT 749
Qy      985  GGAGACCTAAATCGAGCTTACACCAAGCTTAAAGAGTGGAGGATTCATCTGTCAAAAG 1044
Db      750  GGAGACCTAAATCGAGTCCACACCAAGCTTCAGAAGTGGAGGATTCATTTGTACCCAG 809
Qy      1045  AATCACTACATTTACCAGGCAATTCATAAATATATATCTCGTTATCGGTTCTCTAGTAA 1104
Db      810  AATTCGCAATTTTACCAGCAATTTCAAGGATAGTATTATATCTATGAAAGCTGTAAAGTAA 869
Qy      1105  ACTCGTGAAGGCGCACACCTCTGTCTTTGAAAACACATGGCACTGGAAACATCTCGCCTT 1164
Db      870  ACTTGTGTAAGGACACACAGGTACTATCATTTGAAAACCTTTGACAATGGGTCTTCTCCATT 929
Qy      1165  GGATCTGTTCTCCATAATA 1183
Db      930  ACACCTTCTTTATATTTA 948

```

```

RESULT 11
AF334603S6
LOCUS   Mus musculus deoxyribonuclease DLAD (Dlad) gene, exon 6 and
DEFINITION complete cds.
ACCESSION AF334608
VERSION   AF334608.1 GI:17046280
KEYWORDS
SEGMENT
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 2337)
REFERENCE Shiohawa,D. and Tanuma,S.-I.
AUTHORS   Isolation and characterization of the DLAD/Dlad genes, which lie
          head-to-head with the genes for urate oxidase
JOURNAL   Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)
MEDLINE   21556924
PUBMED    11700027
REFERENCE 2 (bases 1 to 2337)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Direct Submission
JOURNAL   Submitted (09-JAN-2001) Department of Biochemistry, Science
          University of Tokyo, Ichigaya, Shinjuku-ku, Tokyo 162-0826, Japan
FEATURES
          Location/Qualifiers

```

source	1. .2337	/organism="Mus musculus"	AC127571	Mus musculus chromosome UNK clone RP24-227D4, WORKING DRAFT	197457 bp	DNA	linear	HTG 17-OCT-2002
		/mol_type="genomic DNA"		SEQUENCE, 4 unordered pieces.				
		/strain="C57BL/6"	AC127571					
gene		/db_xref="taxon:10090"	AC127571.2	GI:24080752				
		order(AF334603.1:<1476. .1825,AF334604.1:1. .1374,	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.					
		AF334605.1:1. .1089,AF334606.1:1. .719,AF334607.1:1. .642,	Mus musculus	(house mouse)				
		1. .1182)	Mus musculus					
		/gene="Dlad"	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
mRNA		Join(AF334603.1:<1476. .1588,AF334604.1:665. .842,	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
		AF334605.1:630. .708,AF334606.1:484. .639,						
		AF334607.1:102. .299,486. .1182)						
		/gene="Dlad"						
CDS		/product="deoxyribonuclease Dlad"	McPherson,J.D. and Waterston,R.H.	The sequence of Mus musculus clone				
		Join(AF334603.1:1476. .1588,AF334604.1:665. .842,	Unpublished					
		AF334605.1:630. .708,AF334606.1:484. .639,	McPherson,J.D. and Waterston,R.H.	(bases 1 to 197457)				
		AF334607.1:102. .299,486. .826)	Direct Submission					
		/gene="Dlad"	Submitted (17-JUL-2002)	Genome Sequencing Center, 4444 Forest Park				
		/codon_start=1	Parkway, St. Louis, MO 63108, USA					
		/product="deoxyribonuclease Dlad"	Submitted (17-OCT-2002)	Genome Sequencing Center, 4444 Forest Park				
		/protein_id="AAL34450.1"	On Oct 17, 2002 this sequence version replaced gi:21887002.					
		/db_xref="GI:17046281"						
		/translation="MTAKPRTVLSLLFPALSGVLGTPPLSCENYGEAVDMFLFKL	Center: Washington University Genome Sequencing Center					
		PKTSKASBAGQYLYLDSTROTWNKSLYLINSALGRTQLQHDYETNSINDTAY	Center code: WUGSC					
		LIYNDGVPVSRYSHAKGLVNRTOGFWLIHSVPKPPVGHVYEYTPGSRGGYGO	Web site: http://genome.wustl.edu/gsc/index.shtml					
		TGICITPGYQPEIDRQLLVLPNIYSCFIPSTFHWKLIYMPRMCASSSLKIPVY	Contact: submissions@watson.wustl.edu					
		LAEUHSQGLNFVHFASKSFYDDIFGWIQAOKLKHLLAQTKWKKQELPNSGLPY	----- Project Information -----					
		HVNYSIGVTSKYSFSSRQDSKWCVSIRGSANRWTCIGDLNKLHQALRGGFICT	Center project name: M_BB0227D04					
		KNHYIQAFHKLXLYRGCK"	----- Summary Statistics -----					
exon		486. .1182	Sequencing vector: M13; 0%					
		/gene="Dlad"	Sequencing vector: plasmid; 100%					
		/number=6	Chemistry: Dye-primer ET; 0% of reads					
IGIN			Assembly program: Phrap; version 0.990319					
			Consensus quality: 196117 bases at least Q30					
			Consensus quality: 196344 bases at least Q20					

			* NOTE: This is a 'working draft' sequence. It currently					
			* consists of 4 contigs. The true order of the pieces					
			* is not known and their order in this sequence record is					
			* arbitrary. Gaps between the contigs are represented as					
			* runs of N, but the exact sizes of the gaps are unknown.					
			* This record will be updated with the finished sequence					
			* as soon as it is available and the accession number will					
			* be preserved.					
			1 31865: contig of 31865 bp in length					
			31866 31965: gap of unknown length					
			31966 70958: contig of 38993 bp in length					
			70959 71058: gap of unknown length					
			71059 136518: contig of 65460 bp in length					
			136519 136618: gap of unknown length					
			136619 197457: contig of 60839 bp in length.					
			Location/Qualifiers					
			1. .197457					
			/organism="Mus musculus"					
			/mol_type="genomic DNA"					
			/db_xref="taxon:10090"					
			/chromosome="UNK"					
			/clone="RP24-227D4"					
			1. .31865					
			/note="assembly_name:Contig14"					
			31966. .70958					
			/note="assembly_name:Contig15"					
			71059. .136518					
			/note="assembly_name:Contig16"					
			136619. .197457					
					</			


```

ORIGIN
Query Match 37.6%; Score 460.4; DB 2; Length 197457;
Best Local Similarity 99.8%; Pred. No. 1.7e-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/Note="assembly_name:Contig17"

763 GACATCTTTACAGATGATAGTCTCAAAAGCTTGAAGACACATTTGTTAGCACAAACCTGG 822
Db GACATCTTTACAGATGATAGTCTCAAAAGCTTGAAGACACATTTGTTAGCACAAACCTGG 68166
QY 823 CAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTGCTTACCATGCTTCAACATC 882
Db CAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTGCTTACCATGCTTCAACATC 68226
QY 883 AAGTCCATTGGGGTAACTTCCAGTCTTACTCAGTCTCGCCAGACCATTTCAATGG 942
Db AAGTCCATTGGGGTAACTTCCAGTCTTACTCAGTCTCGCCAGACCATTTCAATGG 68286
QY 943 TGTGTTTCCATTAAGGGCTCCGCAAAATCGCTGACCTGCAATGGAGACCTTAAATCGAAGC 1002
Db TGTGTTTCCATTAAGGGCTCCGCAAAATCGCTGACCTGCAATGGAGACCTTAAATCGAAGC 68346
QY 1003 CTACACACAGCCTTAAGAGTGAGGATTCATCTGTACAAAAGATCACTACATTTACCAG 1062
Db CTACACACAGCCTTAAGAGTGAGGATTCATCTGTACAAAAGATCACTACATTTACCAG 68406
QY 1063 GCATTTCAATAATATATCTCCGTTATGGTCTGTAACTGGAACCTGGTGAAGGCCACA 1122
Db GCATTTCAATAATATATCTCCGTTATGGTCTGTAACTGGAACCTGGTGAAGGCCACA 68466
QY 1123 CCTCTGTCCTTGAACACTGSCACTGGAACATCTCGCTTGGATCTGTTCTCCATAAT 1182
Db CCTCTGTCCTTGAACACTGSCACTGGAACATCTCGCTTGGATCTGTTCTCCATAAT 68526
QY 1183 ATCAAGGCTTCTGAGTGAGACACAACTAGGCTCCATAAAG 1224
Db ATCAAGGCTTCTGAGTGAGACACAACTAGGCTCCATAAAG 68568

RESULT 13
AC141632/c
LOCUS 211003 bp DNA linear HTG 19-MAR-2003
DEFINITION Mus musculus chromosome UNK clone RP24-100G24, WORKING DRAFT
ACCESSION AC141632
VERSION 1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (18-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Mar 19, 2003 this sequence version replaced gi:29029422.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu

----- Project Information -----
Center project name: M_BB0100G24
----- Summary Statistics -----
Sequencing vector: M13; %
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 208115 bases at least Q40
Consensus quality: 208752 bases at least Q30
Consensus quality: 209170 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1455: contig of 1455 bp in length
* 1456: gap of unknown length
* 1556: contig of 5011 bp in length
* 6566: gap of unknown length
* 6567: contig of 3447 bp in length
* 10113: gap of unknown length
* 10213: contig of 9476 bp in length
* 19689: gap of unknown length
* 19790: contig of 11489 bp in length
* 31279: gap of unknown length
* 31279: contig of 19911 bp in length
* 51290: gap of unknown length
* 51290: contig of 30491 bp in length
* 81881: gap of unknown length
* 81881: contig of 42769 bp in length
* 124750: gap of unknown length
* 124850: contig of 39869 bp in length
* 164719: gap of unknown length
* 164819: contig of 46185 bp in length.

FEATURES
Location/Qualifiers
1..211003
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-100G24"
1..1455
/note="assembly_name:Contig31"
1556..6566
/note="assembly_name:Contig32"
6567..10113
/note="assembly_name:Contig33"
10214..19689
/note="assembly_name:Contig34"
19790..31278
/note="assembly_name:Contig35"
31379..51289
/note="assembly_name:Contig36"
51390..81880
/note="assembly_name:Contig37"
81981..124749
/note="assembly_name:Contig38"
124850..164718
/note="assembly_name:Contig39"
164819..211003
/note="assembly_name:Contig40"

ORIGIN
Query Match 37.6%; Score 460.4; DB 2; Length 211003;
Best Local Similarity 99.8%; Pred. No. 1.7e-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



```

misc_feature 32998..76125
/notes="assembly_fragment"
misc_feature 76226..132416
/notes="assembly_fragment"
misc_feature 132517..255925
/notes="assembly_fragment"
ORIGIN
Query Match 37.6%; Score 460.4; DB 2; Length 255925;
Best Local Similarity 99.8%; Pred. No. 1 Re-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 763 GACATCTTTACAGATGGATAGTCAAAAGTTGAAGACACATTGTTAGCACAACCTGG 822
Db 73338 GACATCTTTACAGATGGATAGTCAAAAGTTGAAGACACATTGTTAGCACAACCTGG 73397
QY 823 CAGAAAAGAAAAGAGAGCTCTCTCAAACTGTTCCCTGCTTACCATGCTTCAACATC 882
Db 73398 CAGAAAAGAAAAGAGAGCTCTCTCAAACTGTTCCCTGCTTACCATGCTTCAACATC 73457
QY 883 AGTCCATTGGGGTAATCTTCAAGTCTTACTTCAGTTCTCGCCAAAGACCATTCACAAATGG 942
Db 73458 AGTCCATTGGGGTAATCTTCAAGTCTTACTTCAGTTCTCGCCAAAGACCATTCACAAATGG 73517
QY 943 TGTGTTTCCATAAAGGCTCCGCAATCGCTGACCTGCAATGGAGACCTTAAATCGAAGC 1002
Db 73518 TGTGTTTCCATAAAGGCTCCGCAATCGCTGACCTGCAATGGAGACCTTAAATCGAAGC 73577
QY 1003 CTACACCAAGCCTTAGAGGTGAGATTCATCTGTACAAAGATCACTACATTTACCAG 1062
Db 73578 CTACACCAAGCCTTAGAGGTGAGATTCATCTGTACAAAGATCACTACATTTACCAG 73637
QY 1063 GCATTTTCATAAATATATCTCCGTATGTTCTGTAAAGTAAACTCGGTGAAGGCCACA 1122
Db 73638 GCATTTTCATAAATATATCTCCGTATGTTCTGTAAAGTAAACTCGGTGAAGGCCACA 73697
QY 1123 CCTCTGTCTTGAACACATGCACTGGAACATCTCGCTGGATCTGTCTCCATAAT 1182
Db 73698 CCTCTGTCTTGAACACATGCACTGGAACATCTCGCTGGATCTGTCTCCATAAT 73757
QY 1183 ATCAGGCTTCTGAGTGACACACACTAGCTGCTCCATATAAG 1284
Db 73758 TTACAGGCTTCTGAGTGACACACACTAGCTGCTCCATATAAG 73799

RESULT 15
AC118117
LOCUS
DEFINITION
Rattus norvegicus clone CH230-42M22, WORKING DRAFT SEQUENCE, 2
unorderd pieces.
AC118117
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 231698)
Munzy,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinu,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Poster,P.,

```

Fraser,C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,I., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwokelimeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L., Puato,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 231698)

Worley,K.C.

Direct Submission

Submitted (13-APR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 231698)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23672467.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVOQ

Center clone name: CH230-42M22

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 218837 bases at least Q40

Consensus quality: 220614 bases at least Q30
 Consensus quality: 222122 bases at least Q20
 Estimated insert size: 227015; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 230572: contig of 230572 bp in length
 * 230573 230672: gap of unknown length
 * 230673 231698: contig of 1026 bp in length.

FEATURES
 source
 1. 231698
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-42M22"
 complement (228283..228997)
 /note="clone boundary
 clone_end:Sp6
 site:
 end_sequence:BH261023"

ORIGIN
 Query Match 25.2%; Score 308.2; DB 2; Length 231698;
 Best Local Similarity 80.7%; Pred. No. 8.8e-81;
 Matches 372; Conservative 0; Mismatches 88; Indels 1; Gaps 1;
 QY 763 GACATCTTTACAGATGATAGCTCAAAAGTTGAAGACACATTTGTTAGCACAAACCTGG 822
 Db 136791 GACATCTTTGAGCTGATAGCTCAAAAGCTGAAGTGCAATTCCTGGTAGAATCCTGG 136850
 QY 823 CAGAAAAAGAACAAAGAGCTTCCTTCAAACTGTTCCCTGCTTACCATGTCTACAAATC 882
 Db 136851 CAGCGAAAGAACACAGAGCTTCCTTCAAACTGTTCCCTGCTTACCATGTCTACAAATC 136910
 QY 883 AAGTCCATTTGGGTAACTTCCAACTTACTTCACTTCTGCTGCGCAAGACCATTCCAAATGG 942
 Db 136911 AAGGCCATTCGGGGACCTCTCGAGTCCGACTTCCTTCTCATCAGCAATTCCAAATGG 136970
 QY 943 TGTGTTTCCATAAAGGGCTCCGCAATCGCTGGACCTGCATTTGGAGACCTAAATCGAAGC 1002
 Db 136971 TGTGTTTCCAAAGGACTCTCAGGCTCGCTGGACCTGTATTTGGAGACCTAAATCGGAGC 137030
 QY 1003 CTACACCAAGCTTAAAGGTGGAGGATTCATCTGTAACAAGATCACTACATTTACCAG 1062
 Db 137031 CCACATCAAGCTTGAAGTGGAGGATTTATCTCTCGAGAAATCGGTACATTTACCAG 137090
 QY 1063 GCATTTCAATAATATATCTCGTTATCGGTTCTGTAAGTAACTCGGTGAAGGCCACA 1122
 Db 137091 TCATTTGATAGTTAGTTTCCATTTATGCTCTTAACTAGCTTGGGGAAGGTCCCA 137150
 QY 1123 CCCTCTGCTTGAACACACCTGGCACTGGAACTCTGCTTGGATCTGTTCTCCATAAT 1182
 Db 137151 CCCACTGCCCTTAAAGACACCTGGCTCTGGACATCTTGCCTTGGATCTATCTCCATATG 137210
 QY 1183 ATCAAGGCTTCTGATGAGCAACAGTAGCGTCCCAATAAA 1223
 Db 137211 TTAAAGGTTTCTGAGTG-TCCTAACATCTCATCCAATAAAA 137250

Search completed: October 14, 2004, 03:42:37
 Job time : 4732.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:29:11 | Search time 48.2954 seconds
(without alignments)
1156.899 Million cell updates/sec

Title: US-10-790-589-2
Perfect score: 1927
Sequence: 1 MTAKPLRTVLSLLFFALSGV.....KNHIYIQAFLKLYLYGPK 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 838593 seqs, 157832909 residues

Total number of hits satisfying chosen parameters: 838593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1927	100.0	354	US-10-790-589-2	Sequence 2, Appl
2	1293.5	67.1	357	US-10-790-589-4	Sequence 4, Appl
3	1293.5	67.1	361	US-10-170-205E-23316	Sequence 23316, A
4	803	31.3	378	US-10-940-774-8375	Sequence 8375, Ap
5	596.5	31.0	360	PCT-US04-02974-42	Sequence 42, Appl
6	596.5	31.0	360	PCT-US04-12706-30	Sequence 30, Appl
7	596.5	31.0	360	PCT-US04-12706-45	Sequence 45, Appl
8	596.5	31.0	360	US-10-770-668-42	Sequence 42, Appl
9	596.5	31.0	360	US-10-830-828-30	Sequence 30, Appl
10	596.5	31.0	360	US-10-830-828-45	Sequence 45, Appl
11	596.5	31.0	360	US-10-830-828-45	Sequence 45, Appl
12	559	29.0	361	US-10-170-205E-36038	Sequence 36038, A
13	559	29.0	361	US-60-565-632-8103	Sequence 8103, Ap
14	535	27.8	375	PCT-US04-12706-61	Sequence 61, Appl
15	535	27.8	375	US-10-830-828-61	Sequence 61, Appl
16	462.5	24.0	373	PCT-US04-12706-14	Sequence 14, Appl
17	462.5	24.0	373	PCT-US04-12706-38	Sequence 38, Appl
18	462.5	24.0	373	US-10-830-828-14	Sequence 14, Appl
19	462.5	24.0	373	US-10-830-828-38	Sequence 38, Appl
20	462.5	24.0	378	PCT-US04-12706-16	Sequence 16, Appl
21	462.5	24.0	378	US-10-830-828-16	Sequence 16, Appl
22	101.5	5.3	475	PCT-US04-24424-1936	Sequence 1936, Ap
23	101.5	5.3	475	US-10-793-373-26	Sequence 26, Appl
24	101.5	5.3	475	US-10-821-801-499	Sequence 499, App
25	101.5	5.3	475	US-10-913-553-142	Sequence 142, App
26	101.5	5.3	475	US-10-940-774-6067	Sequence 6067, Ap

27	94	4.9	2165	6	US-10-789-400-13	Sequence 13, Appl
28	94	4.9	2165	6	US-10-876-113-29	Sequence 29, Appl
29	94	4.9	2165	6	US-10-876-113-30	Sequence 30, Appl
30	90	4.7	1445	7	US-60-584-405-181	Sequence 181, App
31	89	4.6	1451	6	US-10-170-205E-28165	Sequence 28165, A
32	88.5	4.6	2771	6	US-10-844-295-70	Sequence 70, Appl
33	88	4.6	447	7	US-60-592-978-12635	Sequence 12635, A
34	87.5	4.5	471	7	US-60-592-978-21555	Sequence 21555, A
35	87.5	4.5	1278	1	PCT-US04-07626-51	Sequence 51, Appl
36	87.5	4.5	1279	6	US-10-170-205E-28049	Sequence 28049, A
37	87	4.5	473	7	US-60-592-978-19070	Sequence 19070, A
38	87	4.5	625	6	US-10-917-503-11266	Sequence 11266, A
39	87	4.5	836	6	US-10-425-115-359074	Sequence 359074, A
40	87	4.5	964	6	US-10-917-503-12355	Sequence 12355, A
41	87	4.5	1045	6	US-10-917-503-11431	Sequence 11431, A
42	87	4.5	1094	6	US-10-821-234-1097	Sequence 1097, Ap
43	87	4.5	1798	6	US-10-170-205E-36203	Sequence 36203, A
44	87	4.5	2613	6	US-10-455-772-530	Sequence 530, App
45	87	4.5	2628	6	US-10-455-772-502	Sequence 502, App

ALIGNMENTS

RESULT 1

US-10-790-589-2
; Sequence 2, Application US/10790589
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-589-2

Query Match	100.0%;	Score 1927;	DB 6;	Length 354;
Best Local Similarity	100.0%;	Pred. No. 1.3e-177;		
Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLLFFALSGVLGTPBISCRNEYGEADVDFIYKLPKRTSKASEAGLQYL	60	
Db	1	MTAKPLRTVLSLLFFALSGVLGTPBISCRNEYGEADVDFIYKLPKRTSKASEAGLQYL	60	
QY	61	YLDSTQWTKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG	120	
Db	61	YLDSTQWTKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG	120	
QY	121	HAKGLLVNRTQGFLLIHSVPKPPVHGVEYPTSGRRYGTGTGCTIFGYSQFEIDFOLL	180	
Db	121	HAKGLLVNRTQGFLLIHSVPKPPVHGVEYPTSGRRYGTGTGCTIFGYSQFEIDFOLL	180	
QY	181	VLQPNISYCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT	240	
Db	181	VLQPNISYCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYT	240	
QY	241	DDIFTGWIAQKLTLLAQTWOKKQELPNSCLPVAHVNIKSIQVTSKSYFSSRODHSK	300	
Db	241	DDIFTGWIAQKLTLLAQTWOKKQELPNSCLPVAHVNIKSIQVTSKSYFSSRODHSK	300	
QY	301	WCVSIKGSANRWTCIGDLNRSIHOALRGGGFTCTKNHIYIQAFLKLYRYGCFCK	354	

```
Db 301 WCVSIKGSANRWTCIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLRLRYGFK 354
; Query Match 67.1%; Score 1293.5; DB 6; Length 357;
; Best Local Similarity 66.4%; Pred. No. 2.1e-116;
; Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

RESULT 2
US-10-790-589-4
; Sequence 4, Application US/10790589
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCES: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match 67.1%; Score 1293.5; DB 6; Length 357;
Best Local Similarity 66.4%; Pred. No. 2.1e-116;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLISLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLOYL 60
DB 1 MWARLLRTSFAALLFLGLFGVLGAATISCRNEEGKAVDWFIFYKLPKQKESGTGLEYL 60
QY 61 YLDSTRTQWNKSLYLINSTRSALGRTHLOHYDTHNS--TNDTAYLIYNDGVPGSVNYSRQY 119
DB 61 YLDSTRTSRSKSEQLMNDTKSVLGRTHLOHYDTHNS--TNDTAYLIYNDGVPKPNYSRKY 120
QY 120 GHAKGLLVNRTQGFVLIHSPKPPV--HGVEYPTSGRRYGQIGCITFGYSOFEEIDF 177
DB 121 GHYKGLLVNRTQGFVLIHSPKPPV--HGVEYPTSGRRYGQIGCITFGYNOYEIDS 180
QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCASSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
DB 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 240
QY 238 FYTDDIFTGWIATQKLTLLAQTWKKQKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
DB 241 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 300
QY 298 HSKWCVSIKGSANRWTCIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLRLRYGFK 354
DB 301 HAKWCISQKGTKNRWTCIGDLNRSLSHQAALRGGGFICTQNWQIYQAFGLVLYESCK 357

RESULT 3
US-10-170-205E-23316
; Sequence 23316, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCES: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23316
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-23316
```

```
Query Match 67.1%; Score 1293.5; DB 6; Length 361;
Best Local Similarity 66.4%; Pred. No. 2.1e-116;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLISLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLOYL 60
DB 5 MWARLLRTSFAALLFLGLFGVLGAATISCRNEEGKAVDWFIFYKLPKQKESGTGLEYL 64
QY 61 YLDSTRTQWNKSLYLINSTRSALGRTHLOHYDTHNS--TNDTAYLIYNDGVPGSVNYSRQY 119
DB 65 YLDSTRTSRSKSEQLMNDTKSVLGRTHLOHYDTHNS--TNDTAYLIYNDGVPKPNYSRKY 124
QY 120 GHAKGLLVNRTQGFVLIHSPKPPV--HGVEYPTSGRRYGQIGCITFGYSOFEEIDF 177
DB 125 GHYKGLLVNRTQGFVLIHSPKPPV--HGVEYPTSGRRYGQIGCITFGYNOYEIDS 184
QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCASSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
DB 185 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 244
QY 238 FYTDDIFTGWIATQKLTLLAQTWKKQKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
DB 245 SFLLDDIFAAWMAQRLKTHLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 304
QY 298 HSKWCVSIKGSANRWTCIGDLNRSLSHQAALRGGGFICTKNHYIYQAFHKLRLRYGFK 354
DB 305 HAKWCISQKGTKNRWTCIGDLNRSLSHQAALRGGGFICTQNWQIYQAFGLVLYESCK 361

RESULT 4
US-10-940-774-8375
; Sequence 8375, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCES: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8375
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Human
US-10-940-774-8375

Query Match 31.3%; Score 603; DB 6; Length 378;
Best Local Similarity 37.6%; Pred. No. 1.1e-49;
Matches 135; Conservative 59; Mismatches 147; Indels 18; Gaps 8;

QY 5 PLRTVLISLFFALSGVLGTP--HISCRNEYGEAVDWFIFYKLPK--RTSKASERAGLOYL 61
DB 15 PIAMITPLL---LAALLCVPAGALTCTGDSGQPDVDFVYKLPALRGSGGAQRGLQYKY 71
QY 62 LDSTRTQWNKSLYLINSTRSALGRTHLOHYDTHNS--TNDTAYLIYND--GVPGSVNYSRQY 119
DB 72 LDSSGGWRDGRALINSPEGAVGRSLQPLY--RSNTSQAFLAYLNDPQPPSKAQDSSMR 129
QY 120 GHAKGLLVNRTQGFVLIHSPKPPV--HGVEYPTSGRRYGQIGCITFGYSOFEEID 176
DB 130 GHYKGLLVLDHGGFWLVHSPNPPPPASAAAYSWPHSACTYGTLLCVSPFPFAQFKMG 189
QY 177 FQLLVLPQNIYSCFIPSTFHWKLIYMPRMCASSSLKIPVRYLAELHSAQGLNFVHFAKS 236
DB 190 KQLTYTPWVINYQLEIGIFAQEPFDLENVVKGHVSQEPWNSSITLISQAGVQSFPAKF 249
```

```

RESULT 6
PCT-US04-12706-30
; Sequence 30, Application PC/TUS0412706
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: PCT/US04/12706
; CURRENT FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23

```

Query Match	31.0%; Score 596.5; DB 1;	Length 360;
Best Local Similarity	38.0%; Pred. NO. 4.5e-49;	
Matches 132; Conservative	56; Mismatches 144;	
	Indels 15; Gaps	7
OY	17 LSGVLGTP--EISCRNYGSADVDFYKLPK-RTSKASEEAGLQYLILDTSTQTWNKSL	73
DB	6 LAALLCPFAGALTICYGSDGFVDNVVVFVKLPALRIGSCEAAQRGLQYKYLLDESSGWRDGR	65
OY	74 YLINTSRALSARTIQHLYLDTHNTNDTAYILLYND--GVPGSVNYSROYCHAGKLAVWNR	131

```

Db      66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQSKAQDSSMRGHTKGVLILDHD 123
Qy      132 QGFWLIHSPKPPP---VHGVEYPTSGRRYGTGCTITFGYSQPEEIDFOLLVLQPNYIS 188
Db      124 GGFVLHVSVPNFPFPASSAAYSWPHSACTYGTQTLICVSPFPAQSKMGKQLTYYTPWVYN 183
Qy      189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db      184 YOLEGIFAQEPDLENVVKGHVVSQEPWNSSITLTSQAGAVFQSFPAKFSKFGDDLYSGWL 243
Qy      249 AQKLKTHLLAQTQWKKQELPNSCNSLPVHYVNIKSIQVTSKS--YFSSRODHSKWCVSIK 306
Db      244 AALGTNLQVQFWHKTGVLFPNSCSDIQVNLNVNQIAFPAGPAGPSFNSTEDHSKWCVSFK 303
Qy      307 GSANRWTICIGDLNLSHQALRGGGFICTKNHYIYQAFHKLRYLRYGFC 353
Db      304 GP---WTCVGMNRRNQEGEORGGGTLCALPALMKAQFQPLVKNYQPC 347

```

```

RESULT 8
US-10-770-668-42
; Sequence 42, Application US/10770668
; GENERAL INFORMATION:
; APPLICANT: Wright, Susan C.
; APPLICANT: Larick, James W.
; APPLICANT: Nock, Steffen R.
; APPLICANT: Wilson, David S.
; TITLE OF INVENTION: Cell-killing Molecules and Methods of Use Thereof
; FILE REFERENCE: ABSALUS-08602
; CURRENT APPLICATION NUMBER: US/10/770,668
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 42
; LENGTH: 360
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-770-668-42

```

```

Query Match      31.0%; Score 596.5; DB 6; Length 360;
Best Local Similarity 38.0%; Pred. No. 4.5e-49;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

Qy      17 LSGVLGTP--EISCRNEYGEAVDMFIFKLPK-RTSKASEAGLOLYLDSTROTWNKSL 73
Db      6 LALLCVPAGALTCTGDSGQPDVWFVVKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65

Qy      74 YLINSTRSALGRKTLQHLTYDTHNSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNVRT 131
Db      66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQSKAQDSSMRGHTKGVLILDHD 123

Qy      132 QGFWLIHSPKPPP---VHGVEYPTSGRRYGTGCTITFGYSQPEEIDFOLLVLQPNYIS 188
Db      124 GGFVLHVSVPNFPFPASSAAYSWPHSACTYGTQTLICVSPFPAQSKMGKQLTYYTPWVYN 183

Qy      189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db      184 YOLEGIFAQEPDLENVVKGHVVSQEPWNSSITLTSQAGAVFQSFPAKFSKFGDDLYSGWL 243

Qy      249 AQKLKTHLLAQTQWKKQELPNSCNSLPVHYVNIKSIQVTSKS--YFSSRODHSKWCVSIK 306
Db      244 AALGTNLQVQFWHKTGVLFPNSCSDIQVNLNVNQIAFPAGPAGPSFNSTEDHSKWCVSFK 303

Qy      307 GSANRWTICIGDLNLSHQALRGGGFICTKNHYIYQAFHKLRYLRYGFC 353
Db      304 GP---WTCVGMNRRNQEGEORGGGTLCALPALMKAQFQPLVKNYQPC 347

```

```

RESULT 9
US-10-830-828-30
; Sequence 30, Application US/10830828
; GENERAL INFORMATION:

```

```

; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: US/10/830,828
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 360
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-830-828-30

```

```

Query Match      31.0%; Score 596.5; DB 6; Length 360;
Best Local Similarity 38.0%; Pred. No. 4.5e-49;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

Qy      17 LSGVLGTP--EISCRNEYGEAVDMFIFKLPK-RTSKASEAGLOLYLDSTROTWNKSL 73
Db      6 LALLCVPAGALTCTGDSGQPDVWFVVKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65

Qy      74 YLINSTRSALGRKTLQHLTYDTHNSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNVRT 131
Db      66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQSKAQDSSMRGHTKGVLILDHD 123

Qy      132 QGFWLIHSPKPPP---VHGVEYPTSGRRYGTGCTITFGYSQPEEIDFOLLVLQPNYIS 188
Db      124 GGFVLHVSVPNFPFPASSAAYSWPHSACTYGTQTLICVSPFPAQSKMGKQLTYYTPWVYN 183

Qy      189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db      184 YOLEGIFAQEPDLENVVKGHVVSQEPWNSSITLTSQAGAVFQSFPAKFSKFGDDLYSGWL 243

Qy      249 AQKLKTHLLAQTQWKKQELPNSCNSLPVHYVNIKSIQVTSKS--YFSSRODHSKWCVSIK 306
Db      244 AALGTNLQVQFWHKTGVLFPNSCSDIQVNLNVNQIAFPAGPAGPSFNSTEDHSKWCVSFK 303

Qy      307 GSANRWTICIGDLNLSHQALRGGGFICTKNHYIYQAFHKLRYLRYGFC 353
Db      304 GP---WTCVGMNRRNQEGEORGGGTLCALPALMKAQFQPLVKNYQPC 347

```

```

RESULT 10
US-10-830-828-45
; Sequence 45, Application US/10830828
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: US/10/830,828
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 45
; LENGTH: 360
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(360)
; OTHER INFORMATION: DnaeII (Human CRN-6 homolog)

```



```
; TITLE OF INVENTION: Compositions thereof
; FILE REFERENCE: 38-21 (53403) C
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8103
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Diabrotica virgifera
US-60-579-062-8103

Query Match          29.0%; Score 559; DB 7; Length 361;
Best Local Similarity 35.9%; Pred. No. 1.9e-45;
Matches 122; Conservative 65; Mismatches 143; Indels 10; Gaps 9;

Qy 12 LFFALSGVLG-TPEISCRNEYGEAVDFWFIYKLPKRTSKAS-BEAGLYLYLDSTR-OT 68
Db 6 LFFCLYFIAGRTNALSCDENRDVDFVAYKPIQSGRGDIKAGLGHVLTSTNSST 65

Qy 69 WNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQYGHAKGLLYW 128
Db 66 WTFKRSIGDSSLIQNTLHDFY--YNRTK-LSYILYNDEPNHPK-SAAKGHTKGVVIS 121

Qy 129 NRTQGFWLIHVSVPKPPV-HGYEYPTSGRRYGOTGICITFGYSQPEEIDFOLLVLPNIY 187
Db 122 NAAGGYWLIHVSVPKPPVGVNYSFSTGVKYGQTLFVMTMLTNINAVGLQYQNPQHY 181

Qy 188 SCFIPSTPHWKLIVPRMCANSSSLKI PVRYLAELHSAQGLNFVHFPAKSSFYTDIFGTW 247
Db 182 LQOVLSDIKTRVDPDLASAAQNVVKSAPFYVAQIYSKKVAFSTFAKXRFNQELYEDL 241

Qy 248 IAQKLKTHLLAQTKWKKQELPSNCSLPYHVNIKSIGVTSKSY-FSSRDHSHKVCVSIK 306
Db 242 VAPYLKQDLVVEWETPNAGRLHNSGSRVYKVNIOKINMSVNVNFSNTTSDHSHKAVT-S 300

Qy 307 GSANFTWTCIGDLNRSIHQALRGGGFICTKNHYIYQAFHKL 346
Db 301 NSSSSWCVGDINRAQOHLERGCGGTVCQNQKLSITNFHKL 340

RESULT 14
PCT-US04-12706-61
; Sequence 61, Application PC/TUS0412706
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: PCT/US04/12706
; CURRENT FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(375)
; OTHER INFORMATION: NUC-1
PCT-US04-12706-61

Query Match          27.8%; Score 535; DB 1; Length 375;
Best Local Similarity 34.0%; Pred. No. 4.1e-43;
Matches 123; Conservative 59; Mismatches 148; Indels 32; Gaps 8;

Qy 11 SLLFFALSGVLGT-PEISCRNEYGEAVDFWFIYKLPKRTSKASE---EAGLYLYLDSTR 66
Db 7 AVLIFFLLGVSTYAAFSCKDQSGNDVDFWFAVYKMPIEKDDGSGVTGLAGGVAVIYDVNKK 66

Qy 67 Q-TWKNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGS-----112
Db 67 KGTLTSPAKTLDDNDQAIAYTLQYYDKQND-KTIFHWYNDPEPWGSKSTSGIKLEELIS 125

Qy 113 ----VNYSRQ-----YGHAKGLLVNRTQGFWLIHVSVPKPPVHGYEYPTSGRRYGOT 161
Db 126 NRVSNTYTHEDDSTSTAFGHTKGTTFDGTSGVWLVHVSVPKPPVHGYEYPTSGRRYGOT 185

Qy 162 GCITIFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVPRMCANSSSLKIPVRYLAEL 221
Db 162 GCITIFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVPRMCANSSSLKIPVRYLAEL 221

Query Match          27.8%; Score 535; DB 6; Length 375;
Best Local Similarity 34.0%; Pred. No. 4.1e-43;
Matches 123; Conservative 59; Mismatches 148; Indels 32; Gaps 8;

Qy 11 SLLFFALSGVLGT-PEISCRNEYGEAVDFWFIYKLPKRTSKASE---EAGLYLYLDSTR 66
Db 7 AVLIFFLLGVSTYAAFSCKDQSGNDVDFWFAVYKMPIEKDDGSGVTGLAGGVAVIYDVNKK 66

Qy 67 Q-TWKNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGS-----112
Db 67 KGTLTSPAKTLDDNDQAIAYTLQYYDKQND-KTIFHWYNDPEPWGSKSTSGIKLEELIS 125

Qy 113 ----VNYSRQ-----YGHAKGLLVNRTQGFWLIHVSVPKPPVHGYEYPTSGRRYGOT 161
Db 126 NRVSNTYTHEDDSTSTAFGHTKGTTFDGTSGVWLVHVSVPKPPVHGYEYPTSGRRYGOT 185

Qy 162 GCITIFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVPRMCANSSSLKIPVRYLAEL 221
Db 162 GCITIFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVPRMCANSSSLKIPVRYLAEL 221
```

Db	196	MLCMTFKYAQLKSIQTOLFFNRPNIYSSNLPNMAADNADLAKAIAGOYKQGPFSVIE	245
Qy	222	LHSAQGLNFVHFAXSSEYTDIFTGWIAQKLKTHLLACTWOKKQELPSNCSLPYHVYNI	281
Db	246	LETWAGISFTNFAKSEFNADLYDTLVAPTLLKTLVWETW-RRGSEIPLDCKLTYHANDA	304
Qy	282	KSIGVTSKSYFSSRQDHSKWCVSIKGSANRWTCIGDLNRSIHQALRGGGFICTKNHYIQ	341
Db	305	LSIHVGSITAFSYTKDHSKMAHS-ADWTKPWVCIGDINENWTSQYVRGGGTTCISSSFLWK	363
Qy	342	AP	343
Db	364	AY	365

Search completed: October 14, 2004, 00:43:31
 Job time : 50.2954 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 23:51:40 ; Search time 56.7384 Seconds
(without alignments)
1985.252 Million cell updates/sec

Title: US-10-790-589-4
Perfect score: 1933
Sequence: 1 MMARLLRTSFALLFLGLFGV.....QNWQIQAFOGGLVYVESCK 357

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1933	100.0	357	4	Q9NQW3	Q9NQW3 homo sapien
2	1933	100.0	361	4	Q8WZ79	Q8WZ79 homo sapien
3	1297.5	67.1	356	11	Q9QZK9	Q9QZK9 rattus norv
4	1293.5	66.9	354	11	Q9QY48	Q9QY48 mus musculus
5	1290.5	66.8	354	11	Q8C589	Q8C589 mus musculus
6	835	43.2	153	4	Q8WZ80	Q8WZ80 homo sapien
7	762	39.4	375	12	Q90759	Q90759 fowlpox vir
8	560	25.0	350	11	Q9QZK8	Q9QZK8 rattus norv
9	511	26.4	232	12	Q9J5H0	Q9J5H0 fowlpox vir
10	449	23.2	366	5	Q9VED8	Q9VED8 drosophila
11	209	10.8	338	5	Q9BJL8	Q9BJL8 trichinella
12	191.5	9.9	339	5	Q8GQP5	Q8GQP5 trichinella
13	127	6.6	342	5	Q8GQP4	Q8GQP4 trichinella
14	117	6.1	344	5	Q27073	Q27073 trichinella
15	104	5.4	315	5	Q27076	Q27076 trichinella
16	102.5	5.3	698	16	Q823U3	Q823U3 chlamydomphi

17	98.5	5.1	391	16	Q87ZU7	Q87ZU7 pseudomonas
18	97.5	5.0	697	16	O84355	O84355 chlamydia t
19	96.5	5.0	704	16	Q9K1Z6	Q9K1Z6 chlamydia p
20	95.5	4.9	308	16	Q9A122	Q9A122 streptococ
21	95.5	4.9	308	16	Q878E6	Q878E6 streptococ
22	95.5	4.9	328	16	Q8K8C3	Q8K8C3 streptococ
23	95.5	4.9	970	5	Q9NEH2	Q9NEH2 leishmania
24	95	4.9	912	16	Q7VGK4	Q7VGK4 helicobacte
25	94.5	4.9	308	16	Q8P240	Q8P240 streptococ
26	94.5	4.9	349	9	Q9G044	Q9G044 bacterioph
27	94.5	4.9	349	16	Q932C3	Q932C3 staphylococ
28	94.5	4.9	2295	5	Q9TV98	Q9TV98 plasmodium
29	93.5	4.8	305	16	Q8A6C7	Q8A6C7 bacteroides
30	93.5	4.8	394	16	Q9PMP4	Q9PMP4 campylobact
31	93.5	4.8	913	13	Q90XB9	Q90XB9 gallus gall
32	93	4.8	466	13	Q90XC4	Q90XC4 poephila gu
33	93	4.8	526	16	Q9PPR1	Q9PPR1 ureaplasma
34	93	4.8	2472	4	Q14214	Q14214 homo sapien
35	92	4.8	641	16	Q8X5E6	Q8X5E6 escherichia
36	92	4.8	641	16	Q8X3B3	Q8X3B3 escherichia
37	92	4.8	1051	4	Q81XU4	Q81XU4 homo sapien
38	91.5	4.7	276	5	Q9G2V2	Q9G2V2 trichinella
39	91.5	4.7	930	16	Q7VCX6	Q7VCX6 prochloroco
40	91.5	4.7	984	16	Q8A5H5	Q8A5H5 bacteroides
41	91	4.7	395	7	Q8HWE7	Q8HWE7 mus musculu
42	91	4.7	395	7	Q8HWA5	Q8HWA5 mus musculu
43	91	4.7	395	7	Q8HWA4	Q8HWA4 mus musculu
44	91	4.7	433	6	Q8MIL0	Q8MIL0 oryctolagus
45	91	4.7	433	6	Q8MHY7	Q8MHY7 oryctolagus

ALIGNMENTS

RESULT 1	Q9NQW3	PRELIMINARY;	PRT;	357 AA.
ID	Q9NQW3	PRELIMINARY;	PRT;	357 AA.
AC	Q9NQW3			
DT	01-OCT-2000 (TREMREL. 15, Created)			
DT	01-OCT-2000 (TREMREL. 15, Last sequence update)			
DT	01-JUN-2003 (TREMREL. 24, Last annotation update)			
DE	Deoxyribonuclease II beta.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21272509; PubMed=11376952;			
RA	Krisser R.J., Maclea K.S., Park J.P., Eastman A.;			
RT	*The cloning, genomic structure, localization, and expression of human			
RT	deoxyribonuclease Iibeta.;			
RL	Gene 269:205-216(2001).			
DR	EMBL; AF274571; AAF76893.1; ..			
DR	GO; GO:0004531; F:deoxyribonuclease II activity; IEA.			
DR	GO; GO:0006259; P:DNA metabolism; IEA.			
DR	InterPro; IPR004947; DNaase_II.			
DR	Pfam; PF03265; DNaase II; 1.			
SQ	SEQUENCE 357 AA; 41197 MW; 6565D2036BD20955 CRC64;			
Query Match 100.0%; Score 1933; DB 4; Length 357;				
Best Local Similarity 100.0%; Pred. No. 1.1e-171;				
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MMARLLRTSFALLFLGLFGVLTATISCRNEEGKAVDMFTFYKLPKRONKESGTLGYL	60	
Db	1	MMARLLRTSFALLFLGLFGVLTATISCRNEEGKAVDMFTFYKLPKRONKESGTLGYL	60	
QY	61	YLDSTTTSWRKSEOLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVPKPNYSKY	120	
Db	61	YLDSTTTSWRKSEOLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVPKPNYSKY	120	
QY	121	GHTKGLLLNWRVQGFLLIHSIPQFPPIPEGYDPTGRRNGQSGICITFKYNQYEIDS	180	

```
Db 121 GHTKGLLLNVRVQGFLLIHSIPQPPPEEGYDPTGRRNGQSGICITFKXNQYALIS 180
QY 181 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 240
Db 181 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 240
QY 241 SFLLDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 241 SFLLDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIKLSRHSYFSSYQD 300
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 357

RESULT 2
Q8WZ79
ID Q8WZ79 PRELIMINARY; PRT; 361 AA.
AC Q8WZ79;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Endonuclease DLAD.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21556924; PubMed=11700027;
RA Shiokawa D., Tanuma S.I.;
RT "Isolation and Characterization of the DLAD/DIAD Genes, Which Lie
RL Biochem. Biophys. Res. Commun. 288:1119-1128(2001).
DR EMBL; AF334460; AAL34449.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 361 AA; 41712 MW; C6FDD3D3F58F62CAC0 CRC64;

Query Match 100.0%; Score 1933; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e-171; Indels 0; Gaps 0;
Matches 357; Conservative 0; Mismatches 0;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60
Db 5 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 64
QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120
Db 65 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 124
QY 121 GHTKGLLLNVRVQGFLLIHSIPQPPPEEGYDPTGRRNGQSGICITFKXNOYEALIS 180
Db 125 GHTKGLLLNVRVQGFLLIHSIPQPPPEEGYDPTGRRNGQSGICITFKXNOYEALIS 184
QY 181 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 240
Db 185 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 244
QY 241 SFLLDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 245 SFLLDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIKLSRHSYFSSYQD 304
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db 305 HAKWCISQKGTNRWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 361

RESULT 3
```

```
Q9QZK9
ID Q9QZK9 PRELIMINARY; PRT; 356 AA.
AC Q9QZK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease DLAD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=20025354; PubMed=10558878;
RA Tanuma S.I., Shiokawa D.;
RT "Cloning of a cDNA encoding a rat DNase II-like acid DNase.";
RL Biochem. Biophys. Res. Commun. 265:395-399(1999).
DR EMBL; AF178974; AAP13596.1; -.
DR PUR; JC7131; JC7131.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 356 AA; 40472 MW; 5CBC259F42DD6741 CRC64;

Query Match 67.1%; Score 1297.5; DB 11; Length 356;
Best Local Similarity 66.3%; Pred. No. 2e-112;
Matches 236; Conservative 46; Mismatches 73; Indels 1; Gaps 1;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60
Db 1 MTAQPLKAALPLFLFVALSGVLGTPVISCINEDGKAVDWFYFKLPRTSRGGTGMGLDYL 60
QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 119
QY 121 GHTKGLLLNVRVQGFLLIHSIPQPPPEEGYDPTGRRNGQSGICITFKXNOYEALIS 180
Db 120 GHAKGLLLNVRVQGFLLIHSIPQPPPEEGYDPTGRRNGQSGICITFKXNOYEALIS 179
QY 181 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 240
Db 180 QLLVNCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLTLTQSAQOQKFLHFAKSD 239
QY 241 SFLLDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 240 FYTDDIPAAWMAQRLKTHLLTETWQRKQELPNCSLPYHYVNIKAIRGLQSDFPSSHHD 299
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 356
Db 300 HSKWCVSTKDSQARWTCIGDLNRSPhQA FRSGGFICTQNWQIYQAFQGLVLYYESCK 355

RESULT 4
Q9QY48
ID Q9QY48 PRELIMINARY; PRT; 354 AA.
AC Q9QY48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Deoxyribonuclease DLAD.
GN DNASE2B OR DIAD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57Bl/6; TISSUE=Liver;
RX MEDLINE=99428663; PubMed=10497274;
RA Shiokawa D., Tanuma S.I.
```

RT "DLAD, a novel mammalian divalent cation-independent endonuclease with
RT homology to DNase II.";
RL Nucleic Acids Res. 27:4083-4089 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21556924; PubMed=11700027;
RA Shiohawa D., Tanuma S.;
RT "Isolation and Characterization of the DLAD/Dlad Genes, Which Lie
RT Head-to-Head with the Genes for Urate Oxidase.";
RL Biochem. Biophys. Res. Commun. 288:1119-1128 (2001).
DR EMBL; AF128888; AAF05082.1; -.
DR EMBL; AF334608; AAL34450.1; -.
DR EMBL; AF334603; AAL34450.1; JOINED.
DR EMBL; AF334604; AAL34450.1; JOINED.
DR EMBL; AF334605; AAL34450.1; JOINED.
DR EMBL; AF334606; AAL34450.1; JOINED.
DR EMBL; AF334607; AAL34450.1; JOINED.
DR MGD; MGI:1913283; Dnase2b.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0004520; F:endodeoxyribonuclease activity; IDA.
DR InterPro; IPR004947; Dnase_II.
DR Pfam; PF03265; Dnase_II; 1.
SQ SEQUENCE 354 AA; 40793 MW; 40A5DDG6E9AD278A1 CRC64;

Query Match 66.9%; Score 1293.5; DB 11; Length 354;
Best Local Similarity 66.4%; Pred. No. 4.7e-112;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLLRTSFALLFLGVLGAATISCRNEBGAVDWFTFYKLPKRONKESGETCLEYL 60
DB 1 MTAKPLRTVLSLLFALSGLVGTPEISCRNEGEAVDWIFIKLPKRTSKASEEAGLQYL 60

QY 61 YLDSTTRSRWKEQLMNDTKSVLGRTLQOLYAYASKSNNTAYLIYNDGVPKPVNSRY 120
DB 61 YLDSTRTQWNKSLYLINSTRSALGKTLQHLTYDTHNS-TNDTAYLIYNDGVPKPVNSRY 119

QY 121 GHTKGLLWNRVQGWLIHSIPQPIPEGDYDPTGRRNGOSGICITFKNOYEADIS 180
DB 120 GHAKGLLVNRTQGWLIHSVFKPPV--HGVEYPTSGRRYGTGICITFGYSQFEEDF 177

QY 181 QLLVNCNPNVYSCISIPATPHOELIHPOLCTRASSSEIPGRLITLQSAQOGKFLHFAKSD 240
DB 178 QLLVLQPNYISCFIPSTFHWKLIYMPRCMANSGLKIPVRYLAELHSAQGLNFVHFAKSS 237

QY 241 SFLDDIFAAWMAORLTKHLLTETWQRKQELPSNCSLPYHVNINIKALKSRHSYSSYOD 300
DB 238 FYTDDIFTGWTIAQKLTLLAQTWQKKQELPSNCSLPYHVNINIKSIGVTSKSYFSRQD 297

QY 301 HAKWCISQKGTNRTWCIGDLNRSPOAFRSGGFTCTQWQIYQAFQGLVLYESCK 357
DB 298 HSKWCVSIGKSANRTWCIGDLNRSLOALRGGGFTCTKNHYIYQAFHKLRYLRYGCK 354

RESULT 5
Q8C589 PRELIMINARY; PRT; 354 AA.
AC Q8C589;
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Deoxyribonuclease II beta.
GN DNASE2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK079224; BAC37579.1; -.
DR MGD; MGI:1913283; Dnase2b.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0004520; F:endodeoxyribonuclease activity; IDA.
DR InterPro; IPR004947; Dnase_II.
DR Pfam; PF03265; Dnase_II; 1.
SQ SEQUENCE 354 AA; 40765 MW; 40AFB179AD5D428 CRC64;

Query Match 66.8%; Score 1290.5; DB 11; Length 354;
Best Local Similarity 66.1%; Pred. No. 8.9e-112;
Matches 236; Conservative 46; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLLRTSFALLFLGVLGAATISCRNEBGAVDWFTFYKLPKRONKESGETCLEYL 60
DB 1 MTAKPLRTVLSLLFALSGLVGTPEISCRNEGEAVDWIFIKLPKRTSKASEEAGLQYL 60

QY 61 YLDSTTRSRWKEQLMNDTKSVLGRTLQOLYAYASKSNNTAYLIYNDGVPKPVNSRY 120
DB 61 YLDSTRTQWNKSLYLINSTRSALGKTLQHLTYDTHNS-TNDTAYLIYNDGVPKPVNSRY 119

QY 121 GHTKGLLWNRVQGWLIHSIPQPIPEGDYDPTGRRNGOSGICITFKNOYEADIS 180
DB 120 GHAKGLLVNRTQGWLIHSVFKPPV--HGVEYPTSGRRYGTGICITFGYSQFEEDF 177

QY 181 QLLVNCNPNVYSCISIPATPHOELIHPOLCTRASSSEIPGRLITLQSAQOGKFLHFAKSD 240
DB 178 QLLVLQPNYISCFIPSTFHWKLIYMPRCMANSGLKIPVRYLAELHSAQGLNFVHFAKSS 237

QY 241 SFLDDIFAAWMAORLTKHLLTETWQRKQELPSNCSLPYHVNINIKALKSRHSYSSYOD 300
DB 238 FYTDDIFTGWTIAQKLTLLAQTWQKKQELPSNCSLPYHVNINIKSIGVTSKSYFSRQD 297

QY 301 HAKWCISQKGTNRTWCIGDLNRSPOAFRSGGFTCTQWQIYQAFQGLVLYESCK 357
DB 298 HSKWCVSIGKSANRTWCIGDLNRSLOALRGGGFTCTKNHYIYQAFHKLRYLRYGCK 354

RESULT 6
Q8WZ80 PRELIMINARY; PRT; 153 AA.
AC Q8WZ80;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Lung-specific DLAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21555924; PubMed=11700027;
RA Shiohawa D., Tanuma S.;
RT "Isolation and Characterization of the DLAD/Dlad Genes, Which Lie
RT Head-to-Head with the Genes for Urate Oxidase.";
RL Biochem. Biophys. Res. Commun. 288:1119-1128 (2001).
DR EMBL; AF333389; AAL34448.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; Dnase_II.
DR Pfam; PF03265; Dnase_II; 1.
SQ SEQUENCE 153 AA; 17816 MW; ABF4ED379AFA9FCF CRC64;

Query Match 43.2%; Score 835; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.2e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 MPQLCTRASSSEIPGRLITLQSAQOGKFLHFAKSDSFLDDIFAAWMAORLTKHLLTETW 264


```

Db 1 MPOLCTRASSETIPGRLTLTLOAQOQKFLHFAKSDFLDDIPAAWAORLKHLLTETW 60
QY 265 QRKQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTCIGDLNRS 324
Db 61 QRKQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTCIGDLNRS 120
QY 325 PHQAFSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 121 PHQAFSGGFICTQNWQIYQAFQGLVLYESCK 153

RESULT 7
O90759 PRELIMINARY; PRT; 375 AA.
AC O90759;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II (EC 3.1.22.1).
GN CBL1/DNASEB1.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP-9;
RX MEDLINE=98325194; PubMed=9658122;
RA Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;
RT "Fowlpox virus encodes non-essential homologs of cellular alpha-SNAP,
RT PC-1 and an orphan human homolog of a secreted nematode protein.";
RL J. Virol. 72:6742-6751 (1998).
DR EMBL; AJ006408; CAA07012.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
KW Hydrolase.
SQ SEQUENCE 375 AA; 43064 MW; 36C641F74BAD25F6 CRC64;

```

```

Query Match 39.4%; Score 762; DB 12; Length 375;
Best Local Similarity 44.3%; Pred. No. 1.8e-62;
Matches 148; Conservative 64; Mismatches 110; Indels 12; Gaps 7;

QY 26 ISCRNEEGKAVDWFTFFYKLPKRNKESGETGLELYLDSTTRSRKSEQLMNDTKSVLGR 85
Db 47 VNCVNEEGVVDWVFYVYKLPKQ--KLTGKNEVLYIDSNPNKWKRGKVPINSRYSIIGK 104
QY 86 TLOQLYAYASKNNTAYLYNDGVPKPNYSRKVGTGKLLWN--RVQGFMLHSTPQ 143
Db 105 TLPIYDLYDSK--YIYIFPYNDGIPGSKNYSSKVGHGTGKGMWNSDSVTGFVLIHSVPR 162
QY 144 FPPIPEGYDYPPTGRNGGSGICITFKY-NOYEADISQLLVNPNVYSCSIPATFHOEL 202
Db 163 FPPSPVGVNYPYSGYVYGQSMCLINDLYKGLTALDNTLPVNNPVNCSVT---NKQL 219
QY 203 IHMPQLCTRASSETIPGRLTLTLOAQOQKFLHFAKSDFLDDIPAAWAORLKHLLTETW 262
Db 220 NNLVHLCDKNDKNTYTLXNVSRWMSRKGKELTFAKSKYFRHIDMSAWIGPTLESDDLSE 279
QY 263 TWQKRQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTCIGDLN 322
Db 280 TWQKRQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTCIGDLN 337
QY 323 RSPHQAFSGGFICTQNWQIYQAFQGLVLYESCK 356
Db 338 RSPHQAFSGGFICTQNWQIYQAFQGLVLYESCK 371

```

RESULT 8
O9QZK8

```

ID O9QZK8 PRELIMINARY; PRT; 350 AA.
AC O9QZK8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=20025354; PubMed=10558878;
RA Tanuma S.; Shiohara D.;
RT "Cloning of a cDNA encoding a rat DNase II-like acid DNase.";
RL Biochem. Biophys. Res. Commun. 265:395-399 (1999).
DR EMBL; AF178975; AAF13597.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 350 AA; 38175 MW; 6483AF5AC13C5111 CRC64;

Query Match 29.0%; Score 560; DB 11; Length 350;
Best Local Similarity 34.6%; Pred. No. 1.1e-43;
Matches 123; Conservative 62; Mismatches 150; Indels 20; Gaps 9;

QY 11 ALLFLGLGVGLGAATISCRNEEGKAVDWFTFFYKLPKRNKESGE---TGLEYLYLDSTTR 67
Db 6 SLLLAALLMW-PAEALSCYDGRPDVWFVYKLP--ANSQSGDKPKWGLMYKYMDQNSE 62
QY 68 SWKSEQLMNDTKSVLGRITLQQLVEAYASKNNTAYLYNDGVPK--PVNYSRKVGHGTGK 125
Db 63 GWQGVGHIDKQGVGLTLQPLVQ---KNSSQLAFLLYNDQPKSSSAODSSRGHTGK 119
QY 126 LLLMNRVQGFMLHSTPQFP-PIPEGYDYPPTGRNGGSGICITFKYNOYEADISQLV 184
Db 120 VLLDQGGFVHVSVPFRFPSPASSGAYSWPNARTYQTLTLCVSLPSPFPFGIGKQITY 179
QY 185 CNPNVYSCSIPATFHOELIHPQLCTRASSETIPGRLTLTLOAQOQKFLHFAKSDSFLD 244
Db 180 TYPVLYDHKLEGIPFAQLPDLEEVTGKHVLRBPWNSVILTFRAGTTFOSFAKFGKPD 239
QY 245 DIPAAWMAORLKHLLTETWQRKQELPNCSPYHYVNIKAIKLSRHS--YFSSYQDHA 302
Db 240 DLTSGLWLAALGTLNLQVFPNPSGILPSCNSGTHKILDVETGTFPGSGPTFNATEDHS 299
QY 303 KWCISQKGTQNRWTCIGDLNRSPhQAFSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 300 KWCVAPSGP--WVCVGMNRNKRRETHRGGGTGLCTQVPALWKAFFSLV--KACK 348

```

RESULT 9

```

O9J5H0 PRELIMINARY; PRT; 232 AA.
ID O9J5H0;
AC O9J5H0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF FPV032 DNase II.
GN FPV032.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Iu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831 (2000).
RN [2]

```

[illegible]

```

OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ISS 534;
RA Liu M., Li C., Garcia-Reyna P., Fu B., Roman T., Boireau P.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF331159; AAK16519.1; -
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SEQUENCE 338 AA; 37155 MW; 44E0897A32795669 CRC64;

Query Match 10.8%; Score 209; DB 5; Length 338;
Best Local Similarity 21.9%; Pred. No. 5.7e-11;
Matches 76; Conservative 55; Mismatches 176; Indels 40; Gaps 12;

QY 23 AATISCRNEGKAVDWFTFKLPKRONKESGETGLELYLDSTTRSRKSEQLMNDTKSV 82
Db 17 AQVATCKNDANDWYFVYKPNVLSSKILQSGV-----NPAAASRANINGQ--- 65

QY 83 LGRTLQQLYAYASKNNTAYLYNDGVPK--PWNYSRKYGHTKGLLMMNRV--QGFWLI 138
Db 66 AGHSIIRTMAFVUVHQAQINVLAYSDDPNLPPNEKSK---TKGVLLVNNAADEAAAFV 122

QY 139 HSIPQFPPIPEGVDYPTGRRNGQSGICITFKYNOYEAIDSQLVCNPNVYSCSIPATF 198
Db 123 HTVPENFLAY-LNAYSWPFAETPKGHMFLCVSNKAHLNSVSKAIRYOEPYVYANNLPAAI 181

QY 199 ---HQLIHPQLCTRASSEIPGRLLTLTQSAQOGKFLHFAKSDSLDIPAAWMAQRL 255
Db 182 LNQMNELFLINGIDVRVTSFLAHTPATKSVQAVANIQAFGKHSKSFADMYARILNR 241

QY 256 KTHLLTETQKQELPNSCLPHVYVNIKAIKL-----SRHSYFSSYQDHAKWC-ISQK 309
Db 242 AASIM--VNSPADARSKSICKGQHLKQIKTSIQLDGVQVSEAR-----DSAKWALIDGK 293

QY 310 GTRKNRWTICGLNRSFHOAFRSGGFCITONWQIYQAFQGLVLYYESC 356
Db 294 NT-----VCFITNDYTATEKRTPGAACVLENAGVYNAPRTAALNVEAC 336

RESULT 12
Q86QP5 PRELIMINARY; PRT; 339 AA.
AC Q86QP5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Newborn larvae stage-specific DNase II.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ISS534;
RA Liu M.Y., Fu B.Q., Lu Q., Wu X.P., Li C.Y., Wang F., Meng X.P.,
RA Yao C.Y., Niu T.X., Boireau P.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY174186; AAO17050.1; -
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SEQUENCE 339 AA; 37406 MW; 6F318435B11B68A2 CRC64;

Query Match 9.9%; Score 191.5; DB 5; Length 339;
Best Local Similarity 21.7%; Pred. No. 2.4e-09;
Matches 80; Conservative 61; Mismatches 180; Indels 47; Gaps 16;

OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ISS 534;
RA Liu M., Li C., Garcia-Reyna P., Fu B., Roman T., Boireau P.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF331159; AAK16519.1; -
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SEQUENCE 338 AA; 37155 MW; 44E0897A32795669 CRC64;

Query Match 10.8%; Score 209; DB 5; Length 338;
Best Local Similarity 21.9%; Pred. No. 5.7e-11;
Matches 76; Conservative 55; Mismatches 176; Indels 40; Gaps 12;

QY 23 AATISCRNEGKAVDWFTFKLPKRONKESGETGLELYLDSTTRSRKSEQLMNDTKSV 82
Db 17 AQVATCKNDANDWYFVYKPNVLSSKILQSGV-----NPAAASRANINGQ--- 65

QY 83 LGRTLQQLYAYASKNNTAYLYNDGVPK--PWNYSRKYGHTKGLLMMNRV--QGFWLI 138
Db 66 AGHSIIRTMAFVUVHQAQINVLAYSDDPNLPPNEKSK---TKGVLLVNNAADEAAAFV 122

QY 139 HSIPQFPPIPEGVDYPTGRRNGQSGICITFKYNOYEAIDSQLVCNPNVYSCSIPATF 198
Db 123 HTVPENFLAY-LNAYSWPFAETPKGHMFLCVSNKAHLNSVSKAIRYOEPYVYANNLPAAI 181

QY 199 ---HQLIHPQLCTRASSEIPGRLLTLTQSAQOGKFLHFAKSDSLDIPAAWMAQRL 255
Db 182 LNQMNELFLINGIDVRVTSFLAHTPATKSVQAVANIQAFGKHSKSFADMYARILNR 241

QY 256 KTHLLTETQKQELPNSCLPHVYVNIKAIKL-----SRHSYFSSYQDHAKWC-ISQK 309
Db 242 AASIM--VNSPADARSKSICKGQHLKQIKTSIQLDGVQVSEAR-----DSAKWALIDGK 293

QY 310 GTRKNRWTICGLNRSFHOAFRSGGFCITONWQIYQAFQGLVLYYESC 356
Db 294 NT-----VCFITNDYTATEKRTPGAACVLENAGVYNAPRTAALNVEAC 336

RESULT 13
Q86QP4 PRELIMINARY; PRT; 342 AA.
AC Q86QP4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNase II (Fragment).
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ISS534;
RA Liu M.Y., Fu B.Q., Lu Q., Wu X.P., Li C.Y., Wang F., Meng X.P.,
RA Yao C.Y., Niu T.X., Boireau P.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY174187; AAO17051.1; -
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR NON TER 1.
DR SEQUENCE 342 AA; 37273 MW; CLCB12088ADEAF72 CRC64;

Query Match 6.6%; Score 127; DB 5; Length 342;
Best Local Similarity 19.9%; Pred. No. 0.0025;
Matches 68; Conservative 64; Mismatches 152; Indels 58; Gaps 16;

QY 36 VDMFTFKLPKRONKESGETGLELYLDSTTRSRKSEQLMNDTKSVLGRTLQQLYAYA 95
Db 23 VDMFTFKLP-----GKSSSHYL-IPTAATDWTVAANI-DDAQQPIHSTMDCR-YFASG 72

QY 96 SKSNNTAYLYNDGVKPNVSNRYGHTKGLL--WNEVOGFWLTHSIQPQPIPEGVDY 154
Db 73 TKQNTN--ITAYSNPYHPFELPMSPCKGVIMAEQIKGFWLVTAKYFENLA----- 124

QY 155 PPTGRRNGQSGICITFKYNOYEAIDSQLVC-----NPNVYSCSIPATFHQLHNPQ--- 207
Db 125 -----GAVGDLFTHEKTTKEA---AAFLCMSYTDVNLRAVAKIIDIYEQPIVFAQRSA 174

QY 208 -LCTRA--SSSEIPGRLL-----TLQSAQOGKFLH-----FAKSD-SFLDDIFA 248

```

```

Db 175 AQATQAFYDSSETQ-KLVNGLHKYQPTASTSGDGIKTLTPPGVTKIFASAPVAYSSEIYS 233
QY 249 ANMAORLKHLLTETWQRKQELPSNCLPYHYVNIKAIKLSRHSYFSSVQDHAKWCISQ 308
Db 234 NYVVKLQKSMQVYTGTTATVLRKSCAGTLKVENGLVITVTKDTRIPIGQDSARWSVPK 293
QY 309 KGTNRWTCIGDLNRSRPHQAFRSGGFICTQNMWYQAFQGLV 350
Db 294 --SDSDFVCLNTRGTANDAKYGASVACVLKSEAAALFRKMI 333

RESULT 14
ID Q27073 PRELIMINARY; PRT; 344 AA.
AC Q27073;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 43 kDa secreted glycoprotein precursor.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92406752; PubMed=1382055;
RA Vassilatis D.K., Despoimier D., Masek D.E., Polvere R.I., Gold A.M.,
RA van der Ploeg L.H.T.;
RT "Analysis of a 43-kDa glycoprotein from the intracellular parasitic
RT nematode trichinella spiralis.";
RL J. Biol. Chem. 267:18459-18465 (1992).
DR EMBL; M95499; AAA30327.1; -.
DR PIR; A44164; A44164.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006255; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
DR Signal.
FT SIGNAL
FT CHAIN
FT SIGNAL
SQ SEQUENCE 344 AA; 37700 MW; 741D6BE0CF56C28C CRC64;

Query Match 6.1%; Score 117; DB 5; Length 344;
Best Local Similarity 20.7%; Pred. No. 0.022;
Matches 74; Conservative 61; Mismatches 142; Indels 80; Gaps 20;

QY 13 LFLGLFGV-----LGAATSCRNEGKAVDNFTFYKLPKRONKESGETG-LEYLYLDS 64
Db 5 IFLSAFWILHNCILQIHAANCTCRATDD-TEWFLLFK-----PVGLLKAKIISP 53
QY 65 TTRWRKSEQLMN-DTKSVLGRTLQQLYEAVASKNNNTAYLLYNDGVKPKVNYSRKYGHT 123
Db 54 ANAGWANDGANNITDSGHAL---VQTIAEWMPILDDMTALGYSNTPPKSTITSQTS-S 109
QY 124 KGLLWNR--VQGFWLIHSIPQ-FPPIPEGYDYPPTGRRNGQSGICITFKYNYEIDS 180
Db 110 KGIWMFGNETTDGFWLLHTFERAF--NSVAWSPSKFTSGHMAICLSISDNVPLI-- 165
QY 181 QLLVCNPNVSCSPATFHQE-LIHMPLCTRAS-----SSEIPGRLL----- 221
Db 166 -----VPALQYQEWIYFGQVSSEKATEFADLTSLIDGSLPTITPPLMNQOT 212
QY 222 LTTLSAQGQKFLHFAKSDSLDDIPAAWMAQRKTHLLTETWQRKQELPSNCLPYHY 281
Db 213 ITTLNSALST--VYYSKTSRRLMYSGLAKVMVNM--RIWAVDNTLTQTCG--GKI 266
QY 282 YNIKAIKLSRHSYFSSVQDHAKWCISQKGTNRWTCIGDLNRSRPHQAFRSGGFICTQ 338
Db 267 GFVKVVK-SPVITDGTQDRSK-----DKSQNAVIDD---KPVFCFTTNGYSTKQ 312

```

RESULT 15

```

Q27076
ID Q27076 PRELIMINARY; PRT; 315 AA.
AC Q27076;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P49 antigen (Fragment).
GN P49.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91246268; PubMed=2038363;
RA Su X., Prestwood A.K., McGraw R.A.;
RT "Cloning and expression of complementary DNA encoding an antigen of
RT Trichinella spiralis.";
RL Mol. Biochem. Parasitol. 45:331-336 (1991).
DR EMBL; M64242; AAA30328.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
DR NON_TER
FT NON_TER
SQ SEQUENCE 315 AA; 34279 MW; 9AAE2AC93A04CCDE CRC64;

Query Match 5.4%; Score 104; DB 5; Length 315;
Best Local Similarity 20.2%; Pred. No. 0.32;
Matches 66; Conservative 58; Mismatches 130; Indels 72; Gaps 18;

QY 37 DWFTFYKLPKRONKESGETG-LEYLYLDSSTTRSWRKSEQLMN-DTKSVLGRTLQQLYEAY 94
Db 6 EMFLLFK-----PVGLLKAKIISPANAGWANDGANNITDSGHAL---VQTIAEW 52
QY 95 ASKSNNTAVLLYNDGVKPKVNYSRKYGHTKGLLWNR--VQGFWLIHSIPQ-FPPIPEG 151
Db 53 GPILDDMTALGYSNTPPKSTITSQTS-SKGILMFGNETTDGFWLLHTFERAF--NSVA 109
QY 152 YDYPPTGRRNGQSGICITFKYNYEIDSQLLVCNPNVSCSPATFHQE-LIHMPLCT 210
Db 110 WSWPSKFTSGHMAICLSISDNVPLI-----VPALQYQEWIYFGQVS 154
QY 211 RAS-----SSEIPGRLL-----LTTLSAQGQKFLHFAKSDSLDDIPAAWMA 252
Db 155 EKATEFADLTSLIDGSLPTITPPLMNQOTITTLNSALST--VYYSKTSRRLMYSGLA 212
QY 253 QRLKTHLLTETWQRKQELPSNCLPYHYVNIKAIKLSRHSYFSSVQDHAKWCISQKTK 312
Db 213 KVMVNM--RIWAVDNTLTQTCG--GKIGFVKVVK-SPVITDGTQDRSK-----DK 260
QY 313 NRWTCIGDLNRSRPHQAFRSGGFICTQ 338
Db 261 SQNAVIDD---KPVFCFTTNGYSTKQ 283

```

Search completed: October 14, 2004, 00:31:14
Job time : 58.7384 secs

This Page Blank (uspto)